UNIVERSITY OF TROMSØ UIT
The Faculty of Health Sciences
The Department of Medical Biology
The Medical Genetics Department, University Hospital in Northern Norway

# A study of possible genetic causes of inherited breast and ovarian cancer in a Norwegian cancer population

MBI-3911 Masters' thesis in biomedicine

Elisabeth Jarhelle

May 2013

## Acknowledgements

This work was performed at the Medical Genetics department at the University Hospital in Northern Norway (UNN) in Tromsø from August 2012 to May 2013.

I would like to start by thanking my two supervisors Marijke Van Gehlue and Hilde Monica Frostad Riise Stensland for their excellent guidance during my work. In addition, I want to thank them for their patience, knowledge and feedback during the writing process of this thesis. I could not have hoped for better supervisors.

I would also like to thank Jan-Olof Winberg who made it possible for me to work on my thesis at the University Hospital by being a co-supervisor from the department of medical biology, University of Tromsø.

I especially want to thank Vigdis Brox and Elisabeth Kjeldsen Buvang for all their help in the lab, for teaching me and for answering a lot of questions during.

I also would like to thank Ida Mette Rønning for providing me with a selection of patients in regards to the screening for mutations in *PALB2* and Lovise Mæhle at the University hospital in Oslo for providing the PAXgene blood samples for investigation of mRNA from *BRCA1* and *BRCA2*.

I would like to thank all the members of the medical genetics department for a good working environment and for their entertaining lunch and coffee breaks, together with their general good spirits.

Finally, I would like to thank family and friends for their support during all of this, with dinner invitations, coffee breaks and phone calls.

Elisabeth Jarhelle

Tromsø, May 2013

# Index

## Summary

Homologous recombination repair (HRR) is an important repair mechanism, and mutations disrupting the function of this machinery might contribute to cancer formation. Several proteins interact in this mechanism, and the two best known are BRCA1 and BRCA2. Mutations in their corresponding genes *BRCA1* and *BRCA2* are found in 40% of hereditary breast and ovarian cancer cases. However, there are still mutations found in these genes that are classified as variants of unknown clinical significance (VUS). Some of these variants identified in Norwegian individuals were further investigated in this study by PCR amplification of cDNA and subsequent sequencing. In addition the *PALB2* gene was screened for mutations, since the encoded protein is essential in the co-localization of BRCA1 and BRCA2 in the repair machinery and mutations in *PALB2* have been reported to be associated with hereditary breast and ovarian cancer.

In total, 43 patients from a Norwegian cancer population were screened for mutations in *PALB2* by the use of M13 tagged sequencing primers. However, no obvious pathogenic variants were detected. Nineteen individuals with *BRCA1* VUS and 18 individuals with *BRCA2* VUS from families with hereditary breast/ovarian cancer were investigated for the expression of both alleles in lymphocytes as well as for possible effects on RNA splicing. All possible detectable alleles were expressed in lymphocytes and three of the variants, *BRCA1* c.213-5T>A, *BRCA1* c.5434C>G and *BRCA2* c.68-7T>A were shown to influence splicing during mRNA processing. In addition, 4 different normal alternatively spliced transcripts were identified in *BRCA1* and *BRCA2.*

## Abbreviations

53BP1 – p53 binding protein 1

3'ss – 3' splice site

5'ss – 5' splice site

A – adenine

aa – amino acid

Ala – Alanine

AlignGVGD – Align Grantham Variation; Grantham Deviation

Amp – Ampicillin

APRT - adenine phosphoribosyltransferase

ATM – ataxia telangiectasia mutated

BARD1 - BRCA1-associated RING domain

BAP1 – BRCA1 associated protein-1

BASC – BRCA1 associated surveillance complex

BBP – branch point binding protein

BIC – Breast Cancer Information Core

B-ME – β-mercapto ethanol

bp – Base pairs

BRCA1 – breast cancer susceptibility gene 1

BRCA2 – breast cancer susceptibility gene 2

BRCT – BRCA1 C-terminal

BRIP1 – BRCA1 interacting protein C-terminal helicase 1

C – cytosine

C complex – Catalytic complex

CCDC98 – coiled-coil domain-containing protein 98

CDK – cyclin dependent kinase

cDNA – complementary DNA

ChAM – Chromatin-Association Motif

CHK2/CHEK2 – checkpoint kinase 2

CtBP – C-terminal binding protein

Cys – cysteine

dbSNP – Single Nucleotide Polymorphism database

ddNTP – dideoxy nucleotide triphosphate

DMEM – Dulbecco's Modified Eagle's Medium

DMSO - dimethyl sulfoxide

DNA – deoxynucleic acid

dNTP – dioxy nucleotide triphosphate

DSB – double stranded DNA (dsDNA) breaks

dsDNA – double stranded DNA

DSS1 – deleted in split-hand/split-foot syndrome

E complex – Early complex

EDTA – Ethylenediaminetetraacetic acid

ESE – Exonic splicing enhancer

ESS – Exonic splicing silencer

FBS – Fetal bovine serum

G – guanine

gDNA – genomic DNA

Gln – glutamine

Glu – glutamic acid

Gly – glycine

H2AX – Histone 2A member X

HBOC – Hereditary Breast and Ovarian Cancer

HGMDp – Human Gene Mutation Database Professional

Hmw – High molecular weight

hnRNP – heterogeneous nuclear ribonucleoprotein

HR – Homologous recombination

HRR – Homologous recombination repair

Indel – insertion or deletion

ISE – intronic splicing enhancer

ISS – intronic splicing silencer

LB - Luria broth

Lmw – low molecular weight

Lys – lysine

MDC1 – Mediator of DNA Damage protein 1

Met – methionine

Min – minutes

ml – milliliter

µl – micro liter

MLPA – Multiplex Ligation-dependent Probe Amplification

Mre11 – Meiotic recombination 11

MRN – Mre11-RAD50-Nbs1 complex

MUM1 – MUM1 melanoma associated antigen (mutated) 1

NC – normal control

NES – nuclear export signal

NHEJ – Non-homologous end joining

NLS – nuclear localization signal

nt – nucleotides

NTC – no template control

OB – Oligonucleotide binding

PALB2 – Partner and Localizer of BRCA2

PBS – Phosphate buffered saline

PBST – Phosphate buffered saline with tween 20

pcDNA – Plasmid Cytomegalovirus promoter DNA

PCR – Polymerase Chain Reaction

PolyPhen – Polymorphism Phenotyping

PTEN – Phosphatase and Tensin Homolog

PTIP – PAX interacting (with transcription-activation domain) protein 1

qPCR – quantitative PCR

RAD18 – Rec A recombinase 18

RAD51 – Rec A recombinase 51

RAP80 - receptor associated protein 80

RNA – ribonucleic acid

RPA – replication protein A

RING – really interesting new gene

RNF8 – ring finger protein 8, E3 ubiquitin protein ligase

RRM – RNA recognition motif

RS domain – serine/arginine rich domain

RT – room temperature

SCD – Serine Cluster Domain

Ser – Serine

Sdm – site-directed mutagenesis

Sec - seconds

SF1 – splicing factor 1

SHFM1 – split hand/foot malformation (ectrodactyly) type 1

SIFT – Sorting Intolerant From Tolerant

SNP – single nucleotide polymorphism

snRNP - small nuclear ribonucleoprotein

ssDNA – single stranded DNA

SUMO – Small ubiquitin-like modifier

SR protein – serine- and arginine- rich protein

STK1/LKB1 – serine/threonine kinase 11

T – Thymine

TP53 – Tumor protein 53

V - Volt

Val - Valine

VUS – Variant of Unknown clinical Significance

WT – Wilde type

# 1. Introduction

In Norway over 28.000 people were diagnosed with cancer in 2010. Of these, 11.7% of these patients were diagnosed with breast or ovarian cancer (2.852 and 456 diagnosed with breast and ovarian cancer, respectively) [1]. Although most of the cases were sporadic, some had a clear familial heritage and were caused by germline mutations. Germline mutations are passed through generations, and contribute to increased cancer risk in these families.

Although hereditary breast and ovarian cancer (HBOC) phenotype is inherited in an autosomal dominant pattern, germline mutations in one allele in tumor suppressor genes are inherited in a recessive manner. Consequently, it has been difficult to explain inherited cancer development until a "two hit hypothesis" was formulated by Knudson in 1971 [2]. Indeed, in order to develop cancer both the germline mutation (first hit) and the somatic mutation (second hit) have to occur to initiate tumor formation [3]. Accordingly, HBOC is recessive but the phenotype appears to be dominant.

## 1.1. Hereditary Breast and Ovarian Cancer

There are some general rules for classification of cancer cases as hereditary, depending on cancer-type. In Norway, the following are the criteria for referral when suspecting HBOC [4]: (1) Two close relatives with breast or prostate cancer, diagnosed before the age of 55; (2) Three close relatives with breast cancer, regardless of age; (3) Both breast and ovarian cancer, regardless of age; (4) Breast cancer before the age of 40; (5) Ovarian cancer before the age of 65; (6) Male breast cancer.

In 1994, *breast cancer susceptibility gene 1* (*BRCA1*) was discovered as the first breast cancer susceptibility gene [5]. A year later *breast cancer susceptibility gene 2* (BRCA2) was identified [6]. Together, these two genes are responsible for 40 % of hereditary/familial breast and ovarian cancer cases, as can be seen in Figure 1 from [7]. Since then, many pathogenic mutations have been discovered in these two genes. One thousand six hundred and thirty one mutations in *BRCA1* and 1377 mutations in *BRCA2* have been reported in the Human Gene Mutation Database Professional (HGMDp) [8]. Many of these variants are of unknown clinical significance (VUS) and include synonymous, missense, potential splice mutations and in-frame deletions, and pose a considerable challenge for genetic counseling [8, 9].

But far from all HBOC cases are explained by mutations in these two genes. Rare hereditary cancer syndromes like Li-Fraumeni syndrome and Cowden syndrome can cause increased susceptibility for breast and ovarian cancer through mutations in *tumor protein 53* (*TP53*) and *phosphatase and tensin homolog* (*PTEN*), respectively[10, 11]. In addition, mutation in other genes, such as *checkpoint kinase 2* (*CHEK2*), *ataxia telangiectasia mutated* (*ATM*), *serine/threonine kinase 11* (*STK11/LKB1*),

*MSH2/MLH1* and *NBS1* are rare causes of hereditary breast cancer. However, the underlying genetic cause of approximately 50% of familial/hereditary breast and ovarian cancer cases are still unknown [7]. It is believed that the latter group comprises some still unidentified high penetrance genes, or variants at many moderate or low-penetrance loci. A combination of many moderate or low-penetrance genes might together contribute to increased susceptibility of breast and ovarian cancer [7].

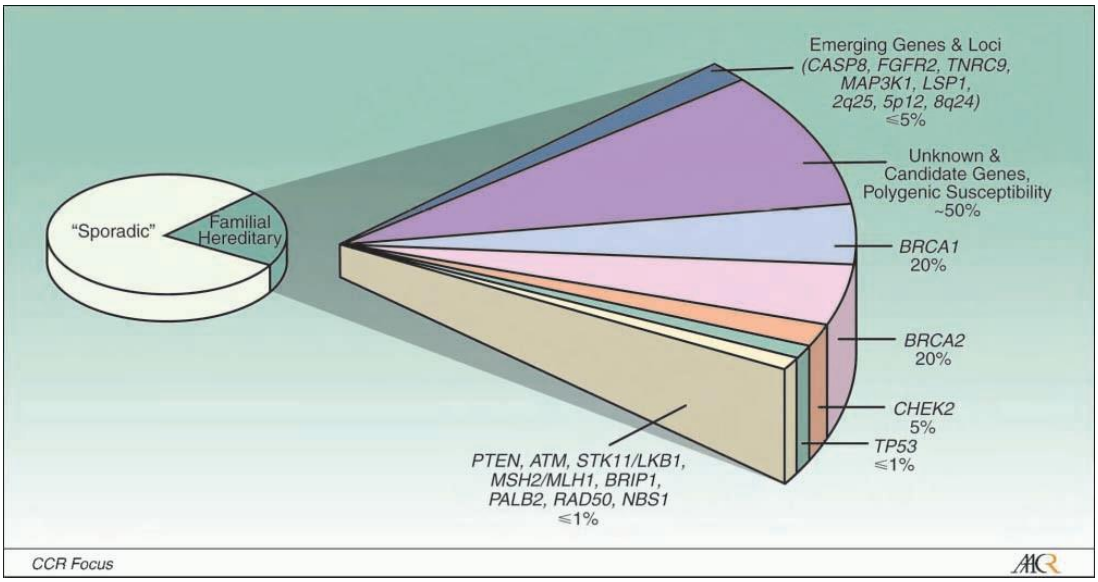



**Figure 1 A diagram showing relative contributions of genes mutated in breast and ovarian cancer [7].**

In recent years, more and more genes with lower penetrance have been documented. *Partner and Localizer of BRCA2* (*PALB2*) has emerged as the 3rd most important gene, after *BRCA1/2* [12]. The *PALB2* gene codes for a protein that plays a key role together with *BRCA1* and *BRCA2* in homologous recombination repair (HRR)[12]. The HRR pathway is responsible for repairing double stranded DNA (dsDNA) breaks (DSB), the most severe form of DNA damage. Fanconi anemia, a hereditary genomic instability disorder has been shown to be caused by bi-allelic mutations in genes that code for proteins involved in HRR [13]. Consequently, the proteins coded by *PALB2*, *BRCA1 interacting protein C-terminal helicase 1* (*BRIP1*), and the *RAD51 homolog (S. cerevisiae)* (*RAD51*) paralogs have been found to increase breast and ovarian cancer susceptibility [14-18].

## 1.2. DNA repair by homologous recombination repair

During their lifetime, our cells undergo many cell divisions. During a cell cycle, there are many things that can go wrong. In most cases these problems are either fixed, or the cell cycle is stopped due to different cell cycle regulation molecules. This haltering of the cell cycle is an important mechanism to

prevent cancer development. Cells that are allowed to replicate although they carry grave genetic mutations, can be precursors to cancer [19].

DNA breaks occur readily, and DSB occur many times per cell per day due to extrinsic factors such as X-rays, UV-light, chemical substances etc., as well as the formation of free radicals [12]. However, the most usual DNA breaks are single stranded DNA breaks. When a ssDNA break occur, the other strand can be used as template [20]. If the DNA break is more severe and causes a DSB, then it is not possible to use the other strand as a template. There are two common DSB repair pathways, where the simplest one is called nonhomologous end-joining (NHEJ). Nonhomologous end-joining simply means that the ends of the DS break are ligated back together, usually resulting in the loss of one or more nucleotides. This mechanism is most common in mammalian somatic cells, where cells usually stop dividing after relatively few cell cycles [20]. Another, and more accurate repair mechanism is homologous recombination repair (HRR) (Figure 2 and Figure 3), which is possible after DNA replication, in the S and G2 phase. Here, the sister chromosome serves as a template and this will accurately restore the DNA sequence [20]. In Figure 2 a simplified view of homologous repair is represented and shows how the sister chromatid functions as a template, while Figure 3 show a more detailed presentation of the assembly of the HRR components.

Homologous repair is a complex repair mechanism, involving several proteins including the aforementioned BRCA1, BRCA2, PALB2, and others. Double stranded breaks are first identified by the MRE11-RAD50-NBS1 (MRN) complex which recruits the ATM protein by phosphorylation [21]. ATM is a cell cycle checkpoint kinase that regulates many downstream proteins. It marks the DSB by phosphorylation of the H2AX chromatin protein as seen in Figure 3. Two of the residues on H2AX are marked and this is recognized by the mediator of DNA-damage checkpoint 1(MDC1) complex. The MDC1 complex consists of MDC1, the MRN complex and the ring finger protein 8, E3 ubiquitin protein ligase (RNF8) [21]. The MDC1 complex subsequently poly-ubiquitylates Lysine (Lys) -63. The poly-ubiquitylation of Lys-63 attracts receptor associated protein (RAP)80, with its double ubiquitin interacting motif, in complex with BRCA1 and ABRAXAS [12]. Abraxas is also designated as coiled-coil domain-containing protein 98 (CCDC98) [22]. The poly-ubiquitylation of Lys-63 also recruits several other proteins, including RAD18 homolog (RAD18), PAX interacting (with transcription-activation domain) protein 1 (PTIP) and p53 binding protein 1 (53BP1). The latter, recruiting a novel DNA damage response protein: MUM1 melanoma associated antigen (mutated) 1 (MUM1/EXPAND1) [21].

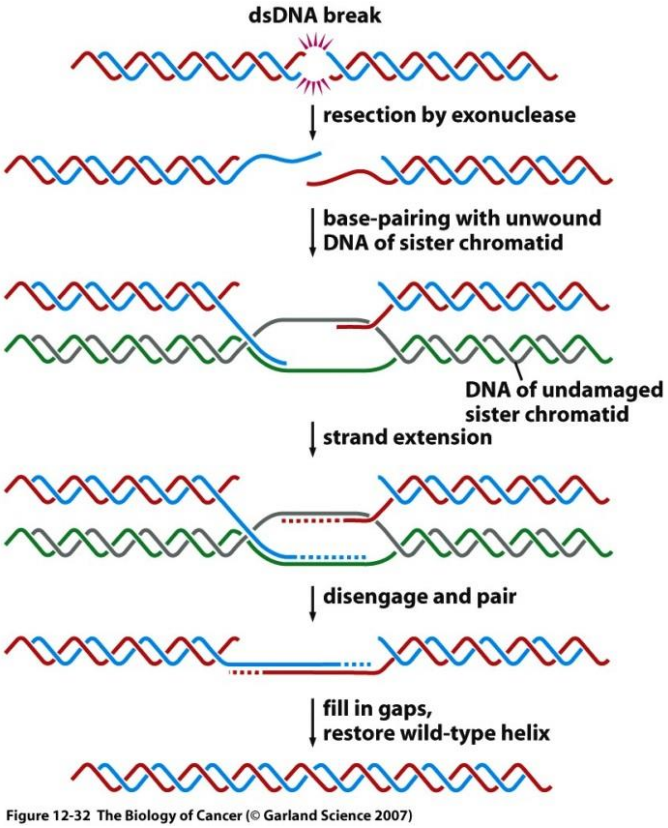



**Figure 2. A simplified presentation of homologous recombination repair in response to a double stranded break. Exonuclease trims the ends at the break, which then can invade the sister chromatid and undergo strand extension. The free ends are thereafter disengaged from the sister chromatid and pair up. The gap is filled, and the helix is restored [23].**

When BRCA1 is recruited, it becomes SUMOylated. SUMO is an ubiquitin-like modifier protein, and attachment of these small proteins functions like a signal mechanism [21]. After SUMOylation of BRCA1, its two trans-activating BRCT-domains successively recruit the MRN complex, C-terminal binding protein (CtBP) and BRIP1. These are recruited to prepare the single stranded extensions that are covered with replication protein A (RPA). BRCA1 binds PALB2, which is in complex with BRCA2 and thereby localizes BRCA2 at the DSB. The BRCA2 protein contains 8 BRC repeats, each binding one RAD51 molecule. RAD51 displaces RPA before performing strand invasion [12]. After strand invasion the sister chromatid is used as template for the extending strands. Thereafter the new strands disengage from the sister chromatid, and pair. The gaps are filled and the helix is restored [23].

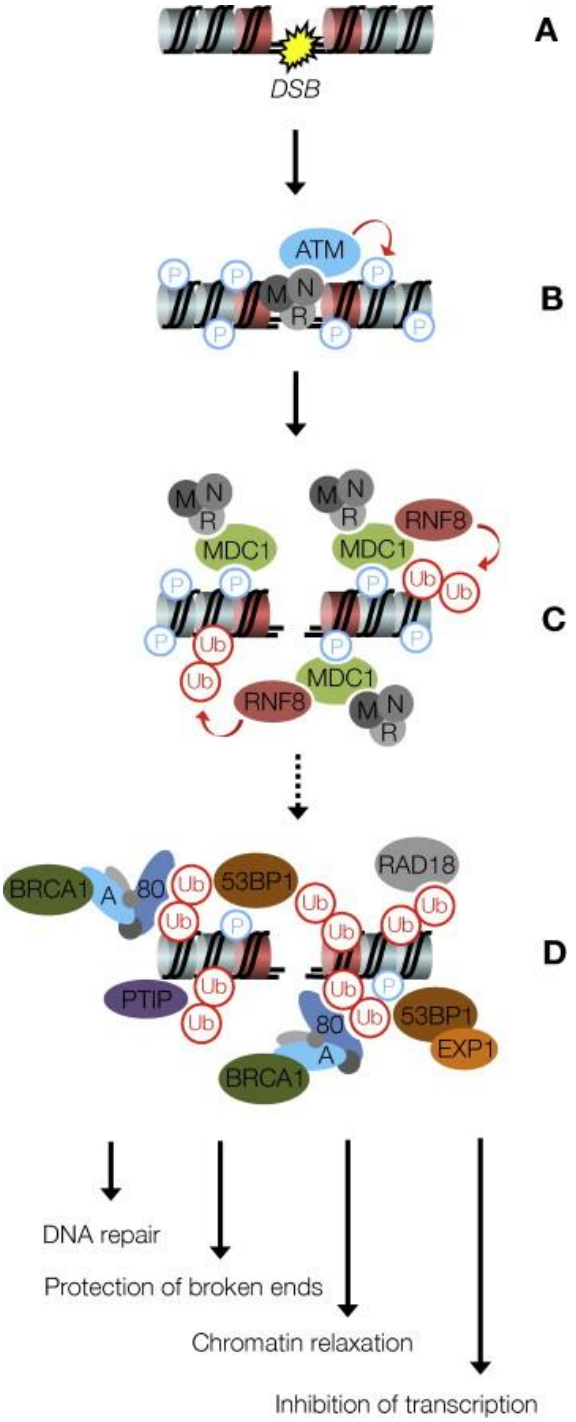



**Figure 3. Foci development in regard to HRR. A) double stranded break. B) The MRN complex recognize the DSB and recruits the ATM kinase. This results in phosphorylation of H2AX by ATM. C) Phosphorylation of H2AX functions as a signal and binding site for MDC1 and partners. This includes the MRN complex and RNF8. RNF8 is a ubiquitin ligase which initiates histone poly-ubiquitylation. D) The poly-ubiquitylation allows a second wave of protein accumulation. This second wave includes 53BP1, BRCA1, RAD18 and PTIP. P: phosphate, M: MRE11, N: NBS1, R: RAD50, Ub: Ubiquitin, A: Abraxas (ABRA1), 80: Rap80, EXP1: EXPAND1 [21].**

### 1.2.1. BRCA1

In 1994, *BRCA1* was discovered as the first breast cancer susceptibility gene [5], but was originally mapped (chromosome 17q21) already in 1990 [24]. The *BRCA1* gene consists of 24 exons and codes for a 220 kD protein consisting of 1,863 amino acids (aa) [22, 25].

At the N-terminal, the BRCA1 protein has a really interesting new gene (RING) domain (aa 8-96), which is a zinc-finger motif with a conserved pattern of cysteine and histidine [26]. The RING domain interacts with BRCA1-associated RING domain (BARD1) protein during cell cycle phase control – G1/S phase control [12, 27]. In addition, BRCA1 has two nuclear localization signal (NLS) at aa 200-300, responsible for BRCA1's nuclear localization [28]; A Serine 988 which requires phosphorylation by checkpoint kinase 2 (CHK2) for promotion of BRCA1-mediated HRR [29]; A coiled-coil domain (aa 1364–1437) shown to be associated with PALB2 [30]; A serine cluster domain (SCD) at aa 1280-1524 that have 10 potentially ATM phosphorylation targets. The phosphorylation of three of these (S1387, S1423 and S1524) is important for the BRCA1-mediated G2/M and S-phase checkpoint activation [31, 32]. At the C-terminal end of BRCA1 (aa 1.646-1.859) there are two BRCA1 C-terminal (BRCT) domains. The BRCT domains each consists of approximately 110 aa. These domains are trans-activating and have many interaction partners, including the MRN complex, CtBP, BRCA1 interacting protein C-terminal helicase 1 (BRIP1) and Abraxas [27, 28]. An overview of these domains can be seen in Figure 4.

BRCA1 is known to be involved in several cellular functions including DNA repair, cell-cycle-checkpoint control, protein ubiquitylation and chromatin remodeling [27].

The first reported mutations in *BRCA1* were truncations; small insertions or deletions (indels), or nonsense mutations leading to an early stop codon. In 1995, the Breast Cancer Information Core (BIC) was established, and subsequently all kinds of mutations have been reported for *BRCA1* [27].

BRCA1 has several different transcripts that result from alternative splicing of exons 1- 11 [33]. It also has two alternative start codons, one resulting in a protein lacking the first 17 aa. In addition there is a pseudogene consisting of only exon 1 and 2 of *BRCA1*, further complicating the investigation of mutations in *BRCA* [33]. According to Ensembl genome browser, there are 32 different transcripts (splice variants) of *BRCA1* [34].

### 1.2.2. BRCA2

The *BRCA2* gene was discovered in 1995 [6], one year after it was mapped to chromosome 13q12-13 [35]. The *BRCA2* gene consists of 27 exons and codes for a protein consisting of 3.418 aa [27, 36].

The BRCA2 protein contains a PALB2 binding domain (aa 18-40), 8 BRC repeats in 80-300 aa spacing (between aa 1009-2083), and two NLS at the C-terminal. Each BRC repeat binds one RAD51 molecule by structural mimicking the RAD51 self-association motif. BRCA2 is thereby responsible for the localization and DNA binding ability of RAD51 [37]. The NLS at the C-terminal binds deleted in split-hand/split-foot syndrome (DSS1) protein [38]. DSS1 is previously known as split hand/foot malformation (ectrodactyly) type 1 (SHFM1) protein [39]. The DSS1 is an evolutionary conserved acidic protein that stabilizes BRCA2 [38].

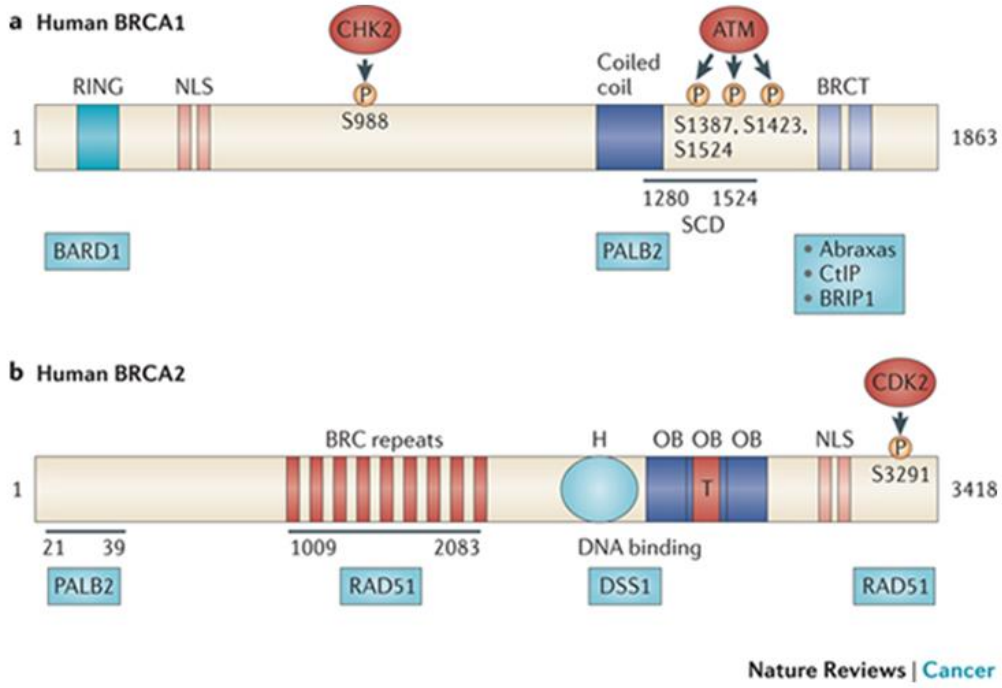



**Figure 4. Domains and binding partners of BRCA1 and BRCA2 [28]. A) BRCA1 contains a RING domain that associates with BARD1 and two NLS at the C-terminal. The central part contains a CHK2 phosphorylation site on S988 [29]. The C-terminal of BRCA1 contains: a coiled-coil domain which associates with PALB2; a SCD that have 10 potential ATM phosphorylation sites; and a BRCT domain that can associate with the MRN complex (not shown), CtBP, BRIP1 and Abraxas. B) The N-terminal of BRCA2 binds PALB2. There are 8 BRC repeats, which each binds a RAD51 molecule. The BRCA2 DNA-binding domain contains a helical domain (H), 3 oligonucleotide binding (OB) folds and a Tower domain (T), which might facilitate BRCA2 binding to both ssDNA and dsDNA. This region also associates with DSS1. At the C-terminal, BRCA2 has 2 NLS and a cyclin dependent kinase (CDK) phosphorylation site, that also binds RAD51 ([28]and references within).**

BRCA2 is best known for its contribution to HRR. But in addition, BRCA2 has several other functions and has been shown to contribute in transcriptional regulation, inhibition of cell proliferation, cytokinesis, mammalian gametogenesis, G2/M checkpoint control, and regulation of centrosome duplication ([38] and references within).

As for *BRCA1*, the first reported mutations in *BRCA2* were truncating mutations. These mutations were easier to classify, in contrast to some of the later detected missense and intronic variants [27]. For *BRCA2*, 6 splice variants are recorded in Ensembl genome browser [40].

### 1.2.3. Partner and localizer of BRCA2 (PALB2)

As the name indicates, partner and localizer of BRCA2 (PALB2) interact with BRCA2. In 2006, PALB2 was identified as a component required for BRCA2-RAD51-loading [41], and is now emerging as the third breast cancer susceptibility gene [12]. *PALB2* is located at 16p12.1 and consists of 13 exons. It codes for a 131 kDa protein consisting of 1,186 aa.

In Figure 5 there is a schematic overview of the domains in PALB2 that were shown to interact with other protein or DNA. As illustrated, the C-terminal end of PALB2 consists of a seven-bladed β-propeller WD40 domain. The WD40 domain is responsible for the BRCA2 binding, and thereby localizing the BRCA2 protein to DSBs. It binds to aa 21-39 at the N-terminal end of BRCA2. PALB2s interaction with BRCA1 is mediated by aa 9-44 in PALB2 and aa 1393-1424 in BRCA1, which comprises the coiled coil region of the two proteins [12]. PALB2 also contains a Chromatin-Association Motif (ChAM), which is required for PALB2 chromatin association [42].

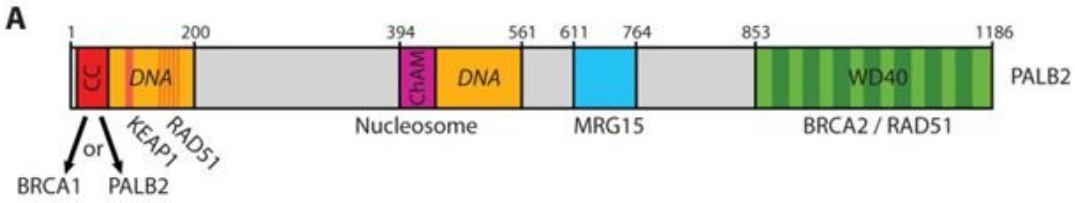



**Figure 5. A schematic presentation of PALB2, with functional domains and corresponding partners (Modified after Rémi Buisson and Jean-Yves Masson, 2012, Nucleic Acids Research [43]. The amino acids numbers are listed above. . CC: Coiled coil domain; ChAM: Chromatin-Association Motif; MRG15: mortality factor 4 like 1.**

PALB2 also interacts with RAD51 directly, promoting filament formation on the ssDNA and strand invasion [43]. The coiled- coil region that interacts with BRCA1 can also bind another PALB2 molecule. Buisson R. and J-Y Masson suggests that the PALB2 self-interaction could be a way to prevent aberrant recombination, and only allow DNA interaction when BRCA1 is active in DSB repair [43]. A more detailed overview of the strand invasion than the previous figure and the PALB2 self-interaction regulation is illustrated in Figure 6 [43].

Most of the identified pathogenic mutations in *PALB2* are nonsense mutations or frameshift leading to truncation. In HGMDp 80 mutations are reported. Mutations are found throughout the gene, indicating that there are no mutational hot-spots [12].

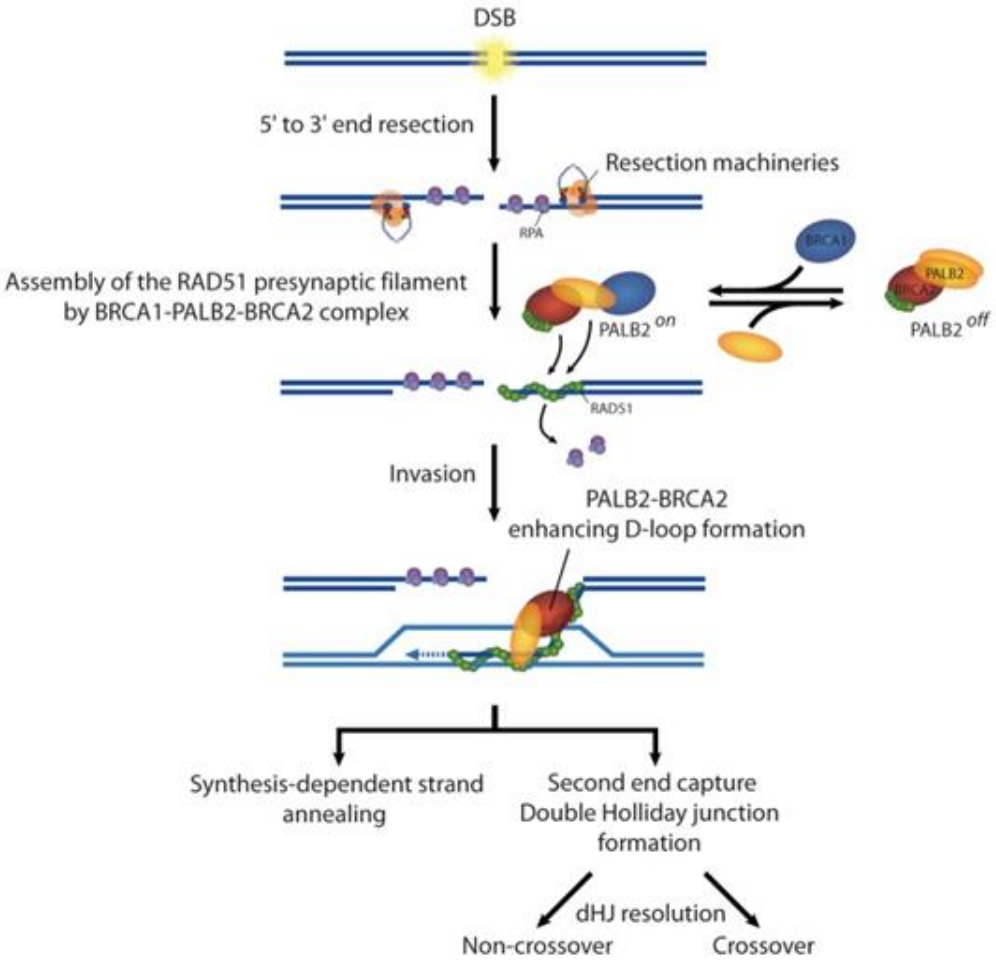



**Figure 6 Model for the role of PALB2 during homologous recombination repair (Modified after Rémi Buisson and Jean-Yves Masson, 2012, Nucleic Acids Research [43]. The BRCA1-PALB2-BRCA2 complex brings RAD51 to the double stranded break, and stimulates RAD51 formation on ssDNA. BRCA2 and PALB2 further stimulate RAD51 during strand invasion.**

## 1.3. mRNA processing – splicing

The average human gene consists of 8.8 exons. The mean exon size is 145 nucleotides (nt), whereas the mean intron length is 3,365 nt. The 5' and 3' UTRs have a mean length of 770 and 300 nt, respectively [44]. Transcription of the DNA produces a ribonucleic acid (RNA) molecule, consisting of the same bases as DNA except that uracil is incorporated instead of thymine. During transcription, the two DNA strands are separated and one of them functions as a template for the RNA to be produced. The transcript is further processed by the splicing machinery, also known as the spliceosome. The spliceosome removes introns, and assembles exons together, see Figure 7. The average mature mRNA is 1340 nt long [45].

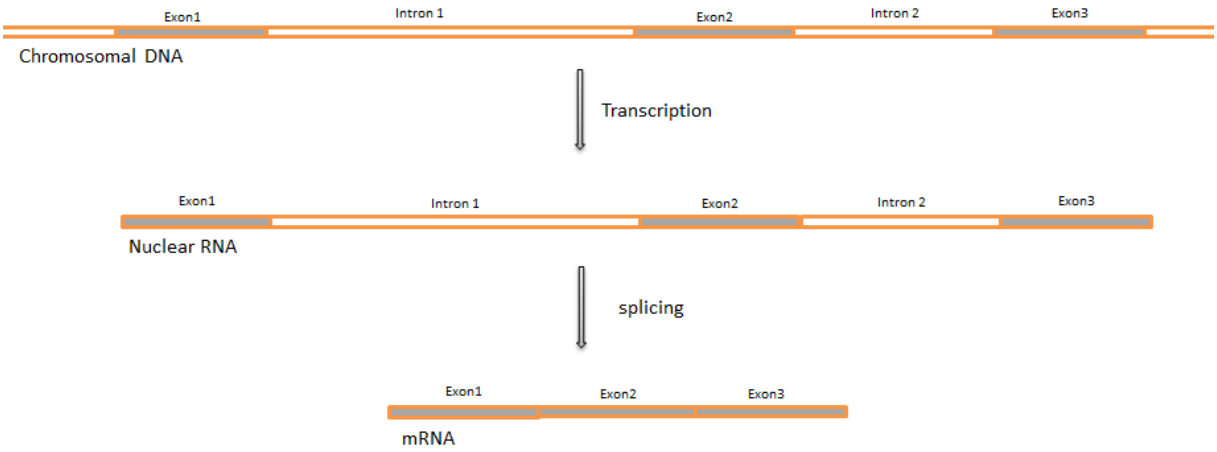



**Figure 7. A simplified figure of mRNA splicing – from DNA to RNA and further to spliced RNA (mRNA).**

### 1.3.1. The Spliceosome

Splicing of pre-mRNA is the removal of introns, to form a mature mRNA consisting of only exon. The splicing machinery consists of several small nuclear ribonucleoprotein particles (snRNPs). These snRNPs, U1, U2, U4, U5 and U6, assemble into pre-mRNA together with other proteins.

The human U1 snRNP consists of the three proteins, U1-70K, U1A and U1C, together with U1 snRNA. The U1 snRNP initiates the first step in splicing by binding the 5' splicing site, and binds a stretch of 4-6 bases at the splice site. To stabilize the binding of snRNA to the 5' splice site, the U1 snRNP is linked to the branch-point binding protein (BBP/SF1) and U2AF through serine- and

arginine-rich (SR) proteins. The BBP and the U2AF (consisting of U2AF65 and U2AF35) binds the branch-point and polypyrimidine tract, respectively. SR proteins consist of one or two N-terminal motifs that recognize RNA and a C-terminal domain rich in Arg/Ser dipeptide repeats. The latter functions as the glue between U1 snRNP, and U2AF and BBP [46].

The bridged U1/U2AF complex is called the commitment complex or E (for early) complex. After formation of the E complex, U2 snRNP basepairs with the branch site and thereby substitutes the BBP in the presence of ATP. The complex is now referred to as the A complex. Thereafter a trimer of U5 and U4/U6 snRNP is recruited to the complex. The now called B1 complex is regarded as the spliceosome, since it contains the required components for splicing. The complex is converted to B2 upon release of the U1 and U4. The release of U1 and U4 allows U6 to come into contact and basepair with the 5' splice site. After a series of rearrangements of the RNA, the complex becomes activated and is now referred to as the C (for catalytic) complex. Subsequently the C complex endures two transesterifications [46]. In the first transesterification, the 5' splice site is cleaved and a lariat with the branch point is formed. The now closest 3' consensus sequence to the 3' side of the lariat is then chosen as the splice site and is the target for the second transesterification [46].

The recruitment of all of the above mentioned components mentioned above are dependent on the sequence of the pre-mRNA. The ends of the introns are highly conserved, but only the first few nucleotides. And the general structure of intron follows the GT-AG rule, which means that the intron start with the dinucleotides GT and ends with AG [46]. Mutations in one of the bases at this position or in the immediately vicinity, or in the poly pyrimidine tract might affect splicing efficiency, and approximately 10% of human hereditable disorders are due to mutations at the splice site junction [47]. Also mutations outside this immediately vicinity might interfere with mRNA splicing, and this fraction might be as high as 25% [47]. These mutations might e.g. abolish one of the splice sites, activate a cryptic site or interfere with recruitment of the splicing machinery [48].

According to Lander *et al.* (2001) [45] approximately 98 % of introns uses the canonical GT-AG pattern. This attracts the U1 and U2 snRNPs as described previously. There are however a minor class of introns that are spliced by the minor spliceosome The minor spliceosome consists of U11 and U12 instead of U1 and U2, respectively, together with the common U5, and $U4_{atac}$ and $U6_{atac}$ snRNAs. These introns are recognized by either AT-AC or GT-AG. The latter corresponds to the sequence recognized by the major spliceosome, but with a different surrounding consensus sequence [46] as seen in Figure 8.

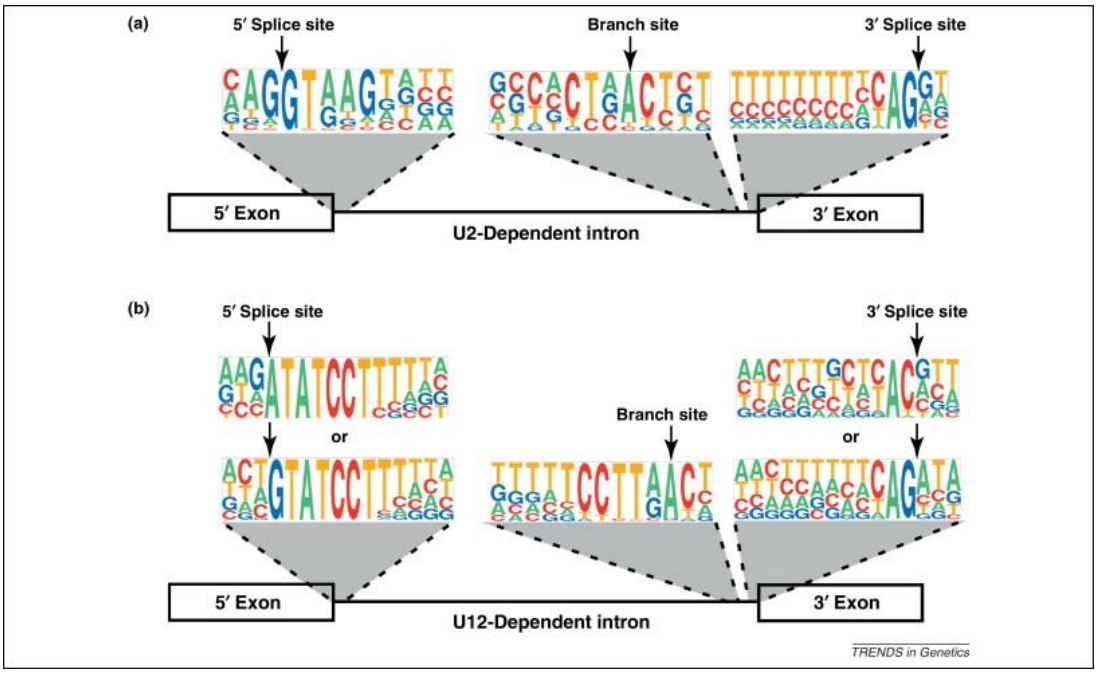



**Figure 8 Splice site consensus sequence for U2-dependent (a) and U12-dependent introns (b), also known as major and minor introns, respectively. This is a graphical representation where the consensus sequence are depicted and the size of the letters represents frequency of the sequences found at that position for all introns [47].**

The consensus-sequences alone are not enough for recognition of exons and introns. Splicing regulatory elements (SRE) are necessary, especially due to the many pseudo splice-sites present in genes [48]. There are 4 different types of regulatory elements: exonic splicing enhancer (ESE); Intronic splicing enhancer (ISE); Exonic splicing silencer (ESS); Intronic splicing silencer (ISS). Together with the consensus sequence, they help in the recognition of exon-intron borders, as well as participate in alternative splicing. Exon-intron boundaries can be recognized either through exon definition or intron definition. Exon definition is mostly used in metazoans, because of the large introns. Exon definition means that SR proteins bind ESE elements. From there, the SR protein recruits the upstream 3' splice site and the downstream 5' splice site localizing snRNPs, as seen in Figure 9 [49].

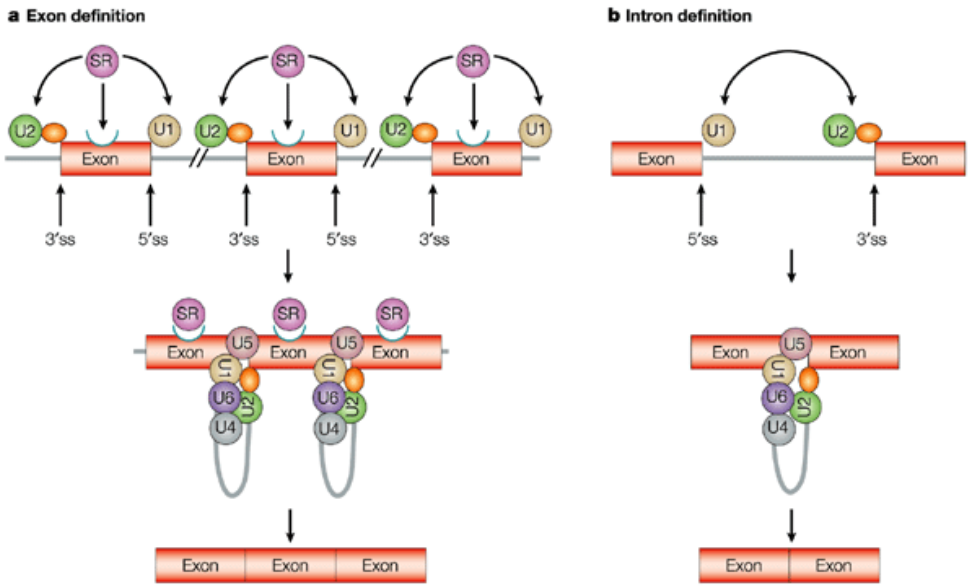



**Figure 9 Exon and intron definition. (a) SR proteins bind to the exonic splicing enhancers (ESE) and recruit U1 to the downstream 5' splice site, and U2AF to the upstream 3' splice site and polypyrimidine tract. U2AF then recruits U2 to the branch site. SR proteins therefore promote cross-exon recognition by recruiting the basal splicing machinery to the flanking ends of the same exon. (b) Intron definition facilitates the binding of the same basal splicing machinery, but at the ends of the same intron instead of exon. SR proteins can also mediate this process [49].**

### 1.3.2. Alternative splicing and disease

Alternative splicing allows for a more divers expression of mRNA. The same gene can code for different proteins, and thereby e.g. have a tissue specific expression. Alternative splicing can therefore regulate localization, enzymatic properties and ligand interactions that the protein might have [50].

In Figure 10, five different alternative splicing options are schematically represented. This includes exon skipping, alternative 5' or 3' splice site, intron retention, and mutually exclusive exons. Alternative spliced exons often deviate more from the consensus splicing sequence and are therefore more dependent on additional splicing factors for recognition. These splicing factors recognizes SREs, and can either bind SR proteins to enhance splicing or heterogeneous nuclear ribonucleoproteins (hnRNP) to inhibit splicing, the latter recognizing ESS and ISS [50].

The SR proteins consist of one or two RNA recognition motifs (RRM) at the N-terminal, and a serine and arginine rich (RS) domain at the C-terminal [51]. The SR protein can form protein-protein interactions through its RS domain, with other proteins containing RS domains. An example of this is the one of the proteins in the U1 snRNP (U1-70K) and in U2AF (U2AF35) [48], as mentioned in 1.3.1. The spliceosome.

.

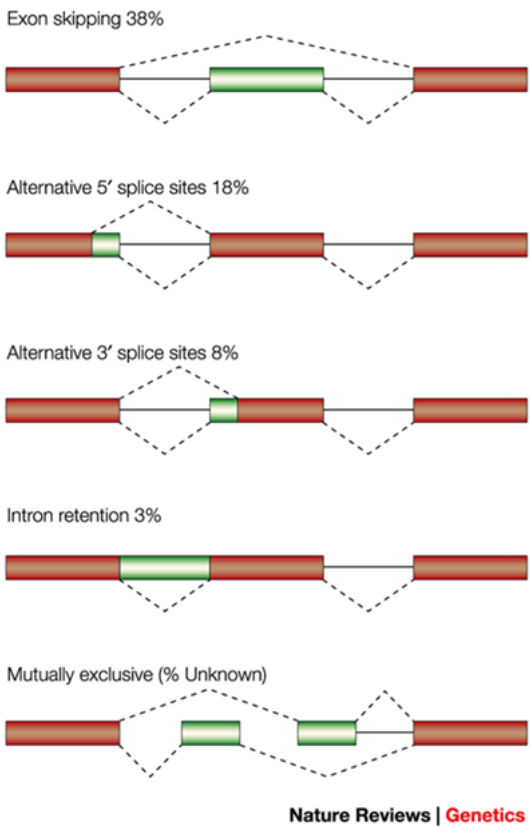



**Figure 10. Five examples of alternative splicing. Constitutive exons are shown in red, while alternative spliced regions are green. Solid lines represents introns, while dashed represents splicing activity. Relative abundance is shown in % of total alternative splicing events [49].**

As mentioned, alternative splicing is normal, but disregulation of splicing might cause disease. Already in 1997, they discovered that silent exonic mutations might be disease causing by affecting splicing of the mRNA [52]. But since then, the focus has shifted to the intronic nucleotides, and as many as 6 % of all reported mutations are now mapped to the GT-AG dinucleotides [48]. Mutations in one of the GT-AG almost always lead to faulty splicing, while the effect of mutations at positions further away from the exon-intron boarder are more difficult to predict due to the fact that their effect might also relies on nucleotides at other positions [53]. There is a direct correlation between amount of faulty splicing and the degree of complementarity between the U1 snRNA and 5'ss [48].

Another example of disregulation of splicing is mutations that create binding sites for hnRNP. This is of special importance for exons with weak splicing signals. Exons with weak splicing signal depend

on the balance between negative and positive SREs and this balance might easily shift due to mutations [48].

## 2. Aim

The aim of this study was to investigate contributing genetic factors in Norwegian families with presumably hereditary breast and ovarian cancer (HBOC). Although about 40% of HBOCs are explained by mutations in *BRCA1* and *BRCA2*,about 60% of the patients familial BC and HBOC for a still unidentified reason. Accordingly, we aimed to investigate mutations if mutations in *PALB2* might be a common cause of cancer in the Norwegian HBOC population. In addition, although sequence variants in *BRCA1* and *BRCA2* have been identified in some individuals, their clinical significance is not always clear. Accordingly, we aimed to study these further by investigating the transcripts of these genes from lymphocytes.

The specific aims of this study:

- To establish an efficient M13 primer based cycle sequencing method
- To screen for *PALB2* mutations in a Norwegian HBOC population
- To investigate the bi-allelic expression and possible effect on pre-mRNA splicing of *BRCA1* or *BRCA2* in blood of individuals with variants of unknown clinical significance in either of these genes.
- To establish a Western blot analysis protocol for future investigation of the consequence of missense mutations in *BRCA1* on the protein level.

# 3. Material and methods

## 3.1. Patient samples

Forty three samples from patients with familial breast and/or ovarian were used in the *PALB2* mutation screening study. In addition, 36 RNA samples of patients carrying variants of unknown clinical significance (VUS) in *BRCA1* or *BRCA2* were used to investigate the influence of these variants on transcription and splicing. All patients have received genetic counseling and signed an informed consent to use their blood samples for genetic analysis.

### 3.1.1. *PALB2*

Forty three patients were included in this study and were chosen based on family history of breast and/or ovarian cancer. Patients included in the screening study of *PALB2* mutations were previously screened for mutations in *BRCA1* and *BRCA2*. No mutations in *BRCA1* and *BRCA2* were found by sequencing the exons and exon/intron boarders of these genes. Thirty five patients were diagnosed with breast cancer, 7 patients had experienced ovarian cancer and one presented both with breast and ovarian cancer. After genetic counseling, Na EDTA blood was collected for DNA isolation. The concentration of the DNA samples was measured using NanoDrop (Thermo Scientific), and subsequently diluted to 50ng/µl. All samples were anonymized and a sample overview can be seen in Table 19 in appendix A.

### 3.1.2. *BRCA1* and *BRCA2*

Thirty six blood samples collected in RNA preserving tubes (PAXgene tubes) from 20 individuals with variants in *BRCA1* (listed in Table 1) and 15 individuals with variants in *BRCA2* (listed in Table 2) of unknown clinical significance were included in this study. These are samples from family members of HBOC-families, but in some of the cases it was not possible to get samples from the affected individual and we therefore examined samples predictively instead. Most of these samples were kindly provided by LM, a cancer geneticist working at Oslo University Hospital. In addition, some samples were collected at the University hospital of Northern Norway (UNN). DNA samples of these individuals were previously screened for *BRCA1* and *BRCA2* mutations. Furtermore, no deletions or duplications were identified in *BRCA1* and *BRCA2* by Multiplex Ligation-dependent Probe Amplification (MLPA). Besides the variants of unknown clinical significance no deleterious

mutations were identified in *BRCA1* and *BRCA2*. RNA samples from these individuals were investigated for the presence of transcripts of both alleles in lymphocytes and their possible impact on splicing of the mRNA.

*In silico* studies were performed to predict the possible effect of the investigated mutations on the BRCA1 and BRCA2 protein.

**Table 1 Variants of unknown clinical significance (VUS) in *BRCA1* investigated in this study.**

| Mutation | Location | Protein | Status | Patient number |
|---|---|---|---|---|
| **c.-33-29delAAAAA** | Intron 1 | | Predictive | 44 |
| **c.1A>G** | Exon 2 | p.Met1? | Predictive | 45 |
| **c.130T>A** | Exon 3 | p.Cys44Ser | Predictive | 46 |
| **c.140G>T** | Exon 4 | p.Cys47Phe | Predictive | 47 |
| **c.213-5T>A** | Intron 4 | | Diagnostic | 48 |
| **c.486G>T** | Exon 7 | p.= (p.Val162Val) | Diagnostic | 49 |
| **c.548-17G>T** | Intron 8 | | Diagnostic | 50 |
| **c.734A>T** | Exon 11 | p.Asp245Val | Diagnostic | 51 |
| **c.1419C>T** | Exon 11 | p.= (p.Asn473Asn) | Diagnostic | 52 |
| **c.1487G>A** | Exon 11 | p.Arg496His | Predictive | 53 |
| **c.2521C>T** | Exon 11 | p.Arg841Trp | Predictive | 54 |
| **c.3418A>G** | Exon 11 | p.Ser1140Gly | Predictive | 55 |
| **c.3707T>G** | Exon 11 | p.Asn1236Lys | Diagnostic | 56 |
| **c.5075A>C** | Exon 18 | p.Asp1692Ala | Predictive | 57 |
| **c.5096G>A** | Exon 18 | p.Arg1699Gln | Predictive | 58 |
| **c.5117G>C** | Exon 18 | p.Gly1706Ala | Predictive | 54 |
| **c.5123C>T** | Exon 18 | p.Ala1708Val | Predictive | 59 |
| **c.5125G>A** | Exon 18 | p.Gly1709Arg | Predictive | 60 |
| **c.5434C>G** | Exon 23 | p.Pro1812Ala | Diagnostic | 61 |
| **c.5513T>G** | Exon 24 | p.Val1838Gly | Predictive | 62 |

**Table 2 Variants of unknown clinical significance (VUS) in *BRCA2* investigated in this study.**

| Mutation | Location | Protein | Status | Patient number |
|---|---|---|---|---|
| **c.40A>G** | Exon 2 | p.Ile14Val | Diagnostic | 63 |
| **c.68-7T>A** | Intron 2 | | Predictive | 64 |
| **c.750G>A** | Exon 9 | p.= (p.Val250Val) | Diagnostic | 65, 76 |
| **c.2680G>A** | Exon 11 | p.Val894Ile | Predictive | 66 |
| **c.3568C>T** | Exon 11 | p.Arg1190Trp | Diagnostic | 67 |
| **c.4068G>A** | Exon 11 | p.= (p.Leu1356Leu) | Diagnostic | 68, 55 |
| **c.4828G>A** | Exon 11 | p.Val1610Met | Diagnostic | 69 |
| **c.5272_5274delAAT** | Exon 11 | p.Asn1758del | Diagnostic | 70 |
| **c.6100C>T** | Exon 11 | p.Arg2034Cys | Diagnostic | 71 |
| **c.6821G>T** | Exon 11 | p.Gly2274Val | Predictive | 72 |
| **c.7301A>C** | Exon 14 | p.Lys2434Thr | Diagnostic | 73 |
| **c.7878G>C** | Exon 17 | p.Trp2626Cys | Diagnostic | 74 |
| **c.8177A>G** | Exon 18 | p.Tyr2726Cys | Diagnostic | 75 |
| **c.8323A>G** | Exon 18 | p.Met2775Val | Predictive | 77 |
| **c.9116C>T** | Exon 23 | p.Pro3039Leu | Diagnostic | 78, 79 |

## 3.2. Molecular methods

### 3.2.1. DNA isolation from blood

DNA was extracted from whole blood, collected in EDTA tubes. Two protocols for DNA isolation were used depending on the DNA extractor used: BioRobot EZ1 or BioRobot M48 (from Qiagen).

The main steps for DNA isolation can be seen in the flowchart provided by the manufacturer shown in Figure 11. In short, lymphocytes are lysed. Subsequently, magnetic particles that bind DNA are added. A magnet is used to contain the DNA in the pipet-tip during subsequent washing. DNA is eluted into new tubes. The whole procedure is performed by the applied BioRobots, and the appropriate reagent kits for each of the extractors were used, following the instructions of the manufacturer (Qiagen).

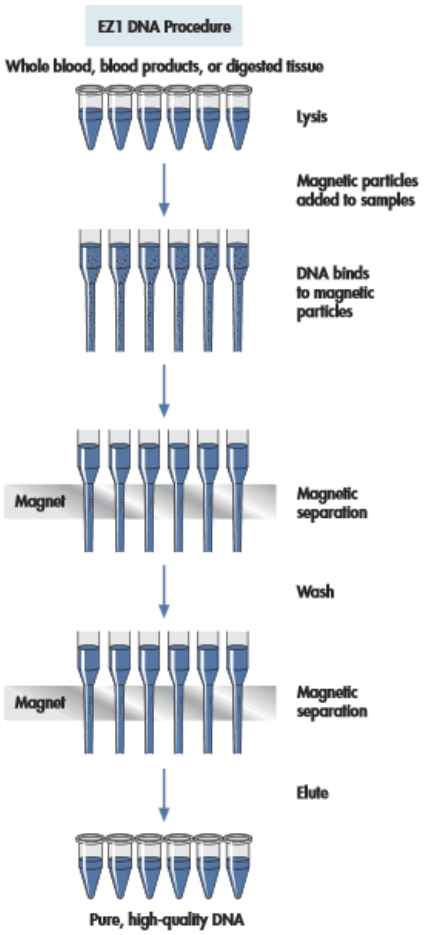



**Figure 11. Main steps in genomic DNA extraction from whole blood using the EZ1 extractor [54].**

### 3.2.2. RNA isolation from blood

Whole blood samples were collected in PAXgene Blood RNA tubes (BRT), which stabilize RNA and are suitable for storage and transport. RNA is stabilized for up to 3 days at 18-25°C and at least 24 months at -20°C or -70°C in these vials [55].

RNA was isolated using QIAGEN PAXgene Blood RNA Kit (PreAnalytiX, Qiagen), according to manufacturer's protocol. The basic steps can be seen in Figure 12. In brief, whole blood stored in PAXgene Blood RNA tubes was first mixed, and subsequently centrifuged (10 minutes at 3000–5000 x g). The pellet was then resuspended and proteinase K was added together with binding buffer. The proteinase K digests residual proteins which were co-precipitated during cell lysis. Ethanol was added to adjust binding conditions. The lysate was transferred to a spin column to remove cell debris. The flow-through supernatant was transferred to a new spin column and added ethanol to adjust binding conditions. Nucleic acids could then bind to the column. The column was washed and treated with DNase before RNA was eluted. After elution the RNA concentration was measured using NanoDrop (Thermo Scientific).

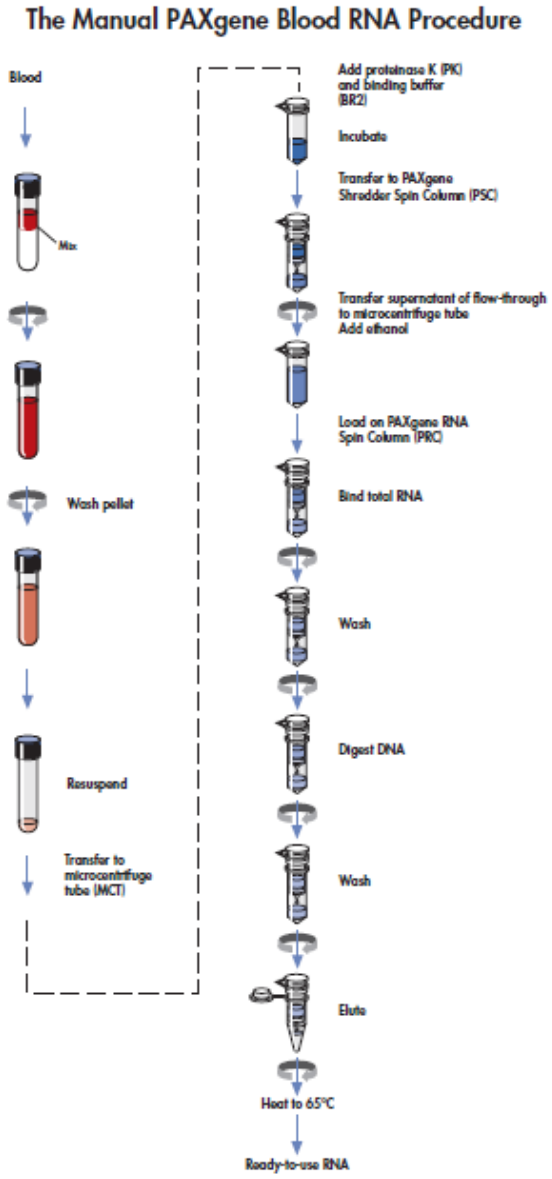



**Figure 12 Basic steps in manual RNA purification from PAXgene Blood RNA kit protocol [55]. Whole blood is first mixed and then centrifuged using a swing-out rotor. The cell pellet is washed, resuspended and transferred to a microcentrifuge tube. The rest of the steps in the procedure are explained in the figure.**

### 3.2.4. cDNA synthesis

Complementary DNA or cDNA is DNA which is copied from RNA using reverse transcriptase. cDNA can be used to indirectly study the gene transcription. cDNA is a more stable molecule to store than RNA. To make cDNA, we used SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen), according to manufacturer's protocol. SuperScript® VILO™ cDNA Synthesis Kit uses random non-specific hexamer primers.

In brief, random oligonucleotide primers anneal to template mRNA, and reverse transcriptase then produces cDNA. Reverse transcriptase uses deoxynucleotides instead of ribonucleotides during incorporation, and thereby produces a DNA copy of the present mRNAs. The newly synthesized DNA strands contain exclusively exons present in the mRNA. Revers transcription was done using MBS 0.2G Thermo hybaid.

The following reaction mixes and PCR machine program were used:

**Table 3 Reaction mixture SuperScript VILO cDNA Synthesis**

| Components | µl per reaction |
|---|---|
| **5X VILO Reaction Mix** | 4 µl |
| **10X Superscript Enzyme Mix** | 2 µl |
| **RNase free water** | 9 µl |
| **RNA** | 5 µl |
| **Total** | **20 µl** |

**Table 4 Modified reaction mixture for samples with RNA concentrations <30ng/µl**

| Components | µl per reaction |
|---|---|
| **5X VILO Reaction Mix** | 4 µl |
| **10X Superscript Enzyme Mix** | 2 µl |
| **RNase free water** | 4 µl |
| **RNA** | 10 µl |
| **Total** | **20 µl** |

cDNA synthesis parameters:

| | | |
|---|---|---|
| Annealing of oligonucleotides | 25 ˚C | 10 min |
| Reverse transcriptase activity | 42 ˚C | 60 min |
| Denaturation of enzyme | 85 ˚C | 5 min |

### 3.2.5. Polymerase chain reaction (PCR)

PCR is used to amplify DNA. A mixture containing polymerase, DNA, nucleotides, primers and a pH-buffer are mixed and then placed in a PCR-machine. The main PCR steps are as follows:

| | |
|---|---|
| Denaturation: | Usually set at 95°C,<br>dsDNA is denatured to ssDNA |
| Annealing: | Approximately 60°C (primer sequence dependent)<br>Primers associate with ssDNA |
| Extension: | Approximately 72°C<br>Polymerase dependent synthesis from the primers 3'-end |

Annealing temperatures set the stringency for primer template recognition, but is dependent on the melting temperature of the primers. Primers are short oligonucleotides normally around 20-22 bases in length. These are reverse complementary to the 5' and 3' end of the sequence to be amplified. Primers used in this study are listed in Table 20, Table 21 and Table 22 in appendix B. For the investigation of VUS's in *BRCA1* and *BRCA2* the name of the corresponding PCR program is listed together with primers. Regarding *PALB2* investigations, BRCA1_2_PCR.HYB PCR program was used for all amplifications except for the amplification of exon 4C product, where the PALB2_4C.HYB PCR program was used.

All PCR reactions (unless otherwise mentioned) were done using JumpStart REDTaq ReadyMix PCR Reaction Mix (Sigma-Aldrich). This reaction mixture is a ready to use mixture where primers, water and DNA should be added prior to the PCR reaction. The reaction mix used throughout this study is as outlined in Table 5, unless otherwise specified.

**Table 5, General reaction mix used, unless otherwise mentioned. The volumes specified are those used for one reaction, and a master mix was usually made.**

| PCR reaction mixture (pr. reaction) | Volume (µl) |
|---|---|
| Primer F (10 pmol/µl) | 0.5 |
| Primer R (10 pmol/µl) | 0.5 |
| Jump start REDTaq | 5.5 |
| Sterile $H_2O$ | 7.5 |
| DNA (50 ng/µl) | 1.0 |
| Total volume: | 15.0 |

### *Touch-down PCR program*

Touch-down is a PCR-program used to avoid unspecific binding of primers during the first cycles of the PCR. Higher annealing temperature in the first cycles increase stringency, and therefore lowers the probability of unspecific binding. Subsequent cycles have lower annealing temperature, until approximately 57°C is reached, where primers bind more easily. Most PCR programs used in this study for amplification are touch-down programs, except programs created after gradient PCR, and the program for the APRT-test. The touch-down programs are listed in appendix A and include the following: (1) BRCA1_2_PCR.HYB, (2) TD63-57_20s.HYB, (3) TD63-57_45s.HYB and (4) TD63-57_2min.HYB.

### *Gradient PCR*

Gradient PCR is the same as regular PCR, only with a gradient of different annealing temperatures. A gradient PCR is used to find the optimal annealing temperature for the primers used. The quality of multiple PCR products representing the multiple primer annealing temperatures used was evaluated after agarose gel electrophoresis. Multiple bands generated by PCR might disturb DNA sequencing. The gradient PCR-program used in this study is listed in appendix A.

### *APRT-test*

After cDNA synthesis, a quality control was performed using the housekeeping gene *Adenine phosphoribosyltransferase* (*APRT*). The *APRT* test is used to asses cDNA quality and purity (regarding genomic DNA). In brief, PCR products are generated with *APRT*-specific primers located in exon 3 and exon 5 (Table 6). Subsequently these products are separated by agarose gel electrophoresis. A fragment of 218 bp is specific for cDNA (NM_000485.2), whereas a 721 bp fragment is generated by amplification of genomic *APRT*-DNA (NC_000016.9). Primers and reaction mixture are specified in Table 6 and Table 7, and the PCR program used is detailed in appendix A.

**Table 6 Primers used in APRT-test, exon 3 and exon 5 specific primers spanning 2 introns (274bp and 229bp), used to amplify cDNA (NM_000485.2) and eventual containing gDNA (NC_000016.9).**

| | Primer sequence (5'-3') | Product size cDNA | Product size gDNA |
|---|---|---|---|
| **hmAPRT1**<br>**hmAPRT2** | GGGGAAGCTGCCAGGCCCCACT<br>GCGAGGTCAGCTCCACCAGGCT | 218 | 721 |

Table 7 reaction mix for APRT-test

| PCR reaction (pr. reaction) | Volum (µl) |
|---|---|
| hmAPRT 1 (20 pmol/µl) | 0.5 |
| hmAPRT 2 (20 pmol/µl) | 0.5 |
| Jump start REDTaq | 12.5 |
| Sterile $H_2O$ | 10.5 |
| cDNA | 1.0 |
| Total volume: | 25.0 |

### 3.2.6. Agarose gelelectrophoresis

Gel electrophoresis is a method used to separate fragments of DNA or RNA based on size. The agarose forms a mesh that allows smaller molecular fragments to move faster than fragments of larger sizes. Movement is due to an electrical field put on the gel, where the negative charged DNA moves from the negative pole towards the positive. Visualization of the DNA bands is possible with UV light by using GelRed, a fluorescent nucleic acid dye.

Procedure:

1. 1.5% or 2% agarose (UltraPure™ from Invitrogen) was dissolved in 1xTBE (Tris Borate EDTA) pH 8.3 buffer and boiled.
2. After cooling to approximately 65°C, the solution was poured into a gel tray. Combs were chosen according to the volume wished to be loaded on the gel. Gels were allowed to solidify for at least 30 minutes.
3. The gel was then placed in an electrophoresis apparatus with 1xTBE buffer and combs were removed
4. Five µl of 1 Kb plus DNA molecular marker (from Invitrogen) was added to the first well (and sometimes last), and PCR products were loaded into the wells.
5. Electrophoresis was run at 100V for 45 minutes (60 minutes for 2.0% agarose gel).
6. Gels were then stained with GelRed (VWR) for 10 minutes and subsequently rinsed in water.
7. DNA was visualized with UV light (BioDoc-It from UVP).

### 3.2.7. Purification of PCR-products

Prior to use of the PCR products in sequencing reactions the PCR products were treated with Illustra ExoStar 1-Step from GE Healthcare Life Sciences. Illustra ExoStar 1-Step contains a mixture of Alkaline Phosphatase and Exonuclease 1. Together they prevent unincorporated primers and nucleotides from the amplification reactions to be used during sequencing reactions.

One µl Illustra ExoStar was used per 12 µl reaction.

Exo-star parameters:

| | | |
|---|---|---|
| Activation | 37°C | 1 h |
| Inactivation | 85°C | 15 min |
| Soak | 4°C | ∞ |

### 3.2.8. Sanger cycle sequencing

Sanger sequencing uses dideoxy nucleotides (ddNTP) for sequence termination. The sequencing reaction mixture contains both deoxy nucleotides (dNTP) and fluorescently labeled ddNTP. DNA strand synthesis is terminated when a ddNTP is incorporated instead of dNTP. Strand termination is due to the lack the 3' hydroxyl group which is required for the formation of phosphodiester bond between the nucleotides. This DNA synthesis results in DNA strands of different sizes, with a fluorescently labeled nucleotide at the 3' end. The strands are separated base on size by the use of capillary electrophoresis. When the fluorophores pass the laser they absorb light which is subsequently emitted and registered by a CCD camera. According to the wavelength of the emitted light, bases are called and registered. The end result is an electropherogram with differently colored peaks, specific to the amplified DNA sequence.

Purified PCR products were sequenced using BigDye v.3.0. from Life Technologies, and cycle parameters are listed below ("Cycle sequencing parameters") together with the corresponding reaction mix (Table 8).

Cycle sequencing parameters:

| | | | |
|---|---|---|---|
| Denaturation | 96°C | 10 sec | 30 cycles |
| Annealing | 50°C | 5 sec | |
| Extension | 60°C | 4 min | |
| Hold | 12°C | ∞ | |

Table 8 Sequencing reaction mix.

| Sequencing reaction (pr. reaction) | Volume (µl) |
|---|---|
| **Primer F or R (10 pmol/µl)** | 0.5 |
| **BigDye v.3.1** | 0.5 |
| **5xBuffer for BigDye v.3.1** | 3.0 |
| **Sterile $H_2O$** | 15.0 |
| **PCR-product (ExoStar -treated)** | 1.0 |
| **Total volume:** | **20.0** |

### 3.2.9. M13 based Sanger cycle sequencing

The primers used for sequencing *PALB2* had an M13 reverse complement sequence attached at the 5' end. This enabled us to use two M13 primers for sequencing various exons of *PALB2*. Sequencing primers are reverse complementary to the M13 sequence incorporated in the PCR products.

The sequence attached to all PALB2 exon specific primers were:

**TGTAAAACGACGGCCAGT** in forward primers

**CAGGAAACAGCTATGACC** in revers primers

PALB2 exon specific primers are listed in Table 20 in appendix B.

Table 9. M13 sequencing primers used in M13-experiment.

| Name | DNA sequence 5'-3' | Supplier | Purification procedure |
|---|---|---|---|
| **M13 F Complement**<br>**M13 R Complement** | ACATTTTGCTGCCGGTCA<br>GTCCTTTGTCGATACTGG | Sigma-Aldrich | Desalting |
| **M13F**<br>**M13R** | ACATTTTGCTGCCGGTCA<br>GTCCTTTGTCGATACTGG | Life Technologies | Desalted and purified |

Two different primers were tested, and are listed in Table 9. Both primer sets have the same M13 sequences, which are the reverse complementary sequences to the primers used for PALB2 PCR (specific *PALB2* primers are listed in Table 20 in appendix B). For amplification of PCR products two different PCR programs were tested: (1) BRCA1_2_PCR.HYB and (2) BRCA1_2_PCR_20s.HYB (listed in appendix A). Sequencing reactions were used as listed in Table 8.

During the M13 based sequencing, two different polymerases were used to generate PCR products: HotStarTaq plus DNA polymerase and RedTaq polymerase. In addition, different template

concentrations and M13 primers from BigDye direct sequencing kit (Life Technologies) were tested (see Table 10), as well as M13 primers from Life Technologies (Table 9).

A 10 µl sequencing mixture using the M13 primers from Life Technologies was also tested, and is listed in Table 11.

**Table 10. Three different template concentrations were used for M13 based sequencing development, using primers from BigDye direct sequencing kit (Life Technologies).**

| Sequencing reaction mix | amount | amount | amount |
|---|---|---|---|
| **Sterile $H_2O$** | 13.0 µl | 12.0 µl | 10.0 µl |
| **5xBigDye buffer** | 4.0 µl | 4.0 µl | 4.0 µl |
| **BigDye mix** | 0.5 µl | 0.5 µl | 0.5 µl |
| **Sequencing primers [? pmol/ µl]** | 1.5 µl | 1.5 µl | 1.5 µl |
| **Template (PCR product)** | 1.0 µl | 2.0 µl | 4.0 µl |
| **Total volume** | **20.0 µl** | **20.0 µl** | **20.0 µl** |

**Table 11. Sequencing mixture used for a 10 µl reaction (M13 primers from Life Technologies).**

| Sequencing reaction (pr. reaction) | Volume (µl) |
|---|---|
| **Primer F or R (0.5 pmol/µl)** | 0.25 |
| **BigDye v.3.1** | 0.25 |
| **5xBuffer for BigDye v.3.1** | 1.5 |
| **Sterile $H_2O$** | 7 |
| **PCR-product (ExoStar -treated)** | 1.0 |
| **Total volume:** | 10.0 |

### 3.2.10. Site-directed mutagenesis

HA-pcDNA BRCA1 Ser1387Ala (Purchased from Addgene) was used to generate HA-pcDNA WT BRCA1 by site-directed mutagenesis (sdm).

For sdm, the QuickChange® XL Site-Directed Mutagenesis Kit (Agilent Technologies) was used. Procedure was done according to manufacturer recommendations. In brief, PCR was performed with primers specifically reversing the mutated nucleotide in the plasmid back to WT nucleotide. The cycling parameters for QuickChange XL were followed (listed under). We used 12 minutes of elongation due to the 12kb plasmid (7.5 bp insertion and 4.5 bp plasmid), since manufacturer recommends 1 min/1kb of plasmid length. Thermal cycler used was Veriti (96 well) from Applied Biosystems.

After the site-directed mutagenesis PCR, template plasmid DNA was removed treating the PCR mixture with DpnI. This is a restriction enzyme recognizes methylated DNA and subsequently cleaves it. DpnI recognizes GATC, where A is methylated. And all template plasmids are methylated since they were synthesized in *Escherichia coli* cells. One µl DpnI is added to the amplification reaction, and the amplification reaction is subsequently incubated at 37˚C for 1 hour.

Subsequently, newly synthesized plasmid was transformated into the XL 10-Gold® Ultracompetent cells. The manufacturers protocol was followed, except for the use of 250 µl S.O.C (Invitrogen) medium instead of 500 µl NZY$^{+}$ broth. In brief, cells were thawed on ice, before 45 µl was transferred to pre-chilled falcon tubes. They were subsequently treated with β-mercapto ethanol (β-ME), to increase transformation efficiency. After 10 minutes, 2 µl DpnI-treated plasmid DNA was added and the mix was left on ice for 30 minutes. The cells were given a 30sec heat-shock (42˚C) and were afterwards returned for 2 min on ice. Cells were subsequently added 250 µl S.O.C. and shaken (225-250 rpm) for 1 hour at 37˚C.

Fifty µl and 200 µl were plated out on agar plates (LB medium) with 100mg/l ampicillin. Bacteria were grown overnight at 37°C. Some of the colonies were selected and tested by PCR for the presence of pcDNA 3-*HA-BRCA1* plasmid. Plasmid DNA was sequenced for verification of the WT sequence of *BRCA1*. Later the entire *BRCA1*-insert was sequenced to confirm that no errors had been incorporated, and primers are listed in Table 22 in Appendix B.

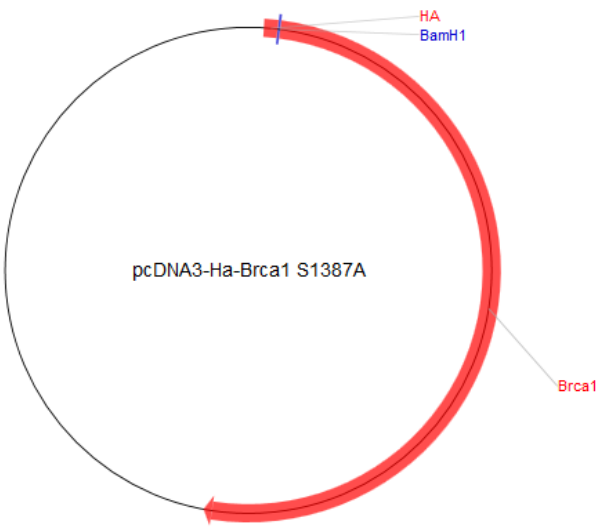



**Figure 13 The plasmid used for site-directed mutagenesis. pcDNA3-Ha-Brca1 S1387A was purchased from AddGene (http://www.addgene.org/32532/)**

**Table 12. Primers used for site-directed mutagenesis of the plasmid back to WT. Nucleotides in bold/underlined corresponds to S1387.**

| | Primer sequence (5'-3') |
|---|---|
| **BRCA1/S1387.F** | TGCTCAGGGCTATCC**TCT**CAGAGTGACATTTTA |
| **BRCA1/S1387.R** | TAAAATGTCACTCTGAGAGGATAGCCCTGAGCA |

Cycling parameters for the QuickChange XL method:

| | | | |
|---|---|---|---|
| Denaturation | 95°C | 1 min | |
| Denaturation | 95°C | 50 sec | |
| Annealing | 60°C | 50 sec | 18 cycles |
| Extension | 68°C | 12 min | |
| Final extension | 68°C | 7 min | |
| Hold | 12°C | | |

### *Plasmid DNA isolation*

Bacteria colonies were cultured overnight at 37°C, 225-250 rpm in LB medium (Appendix B) with ampicillin. Plasmids were purified from the overnight grown bacterial cultures using a mini-prep (Zoppy™ Plasmid Miniprep Kit, from Zymo Research) or midi-prep with NucleoBond® Finalizers (NucleoBond® Xtra Midi from Macherey-Nagel).

Both mini- and midi plasmid preparations were used. The choice depended on amount of plasmid necessary in future studies. For the miniprep used here, lysis buffer could be added directly to the culture. For midiprep the culture medium has to be removed by pelleting cells and removing supernatant. The following steps are basically the same for both procedures.

The cells are lysed under alkaline conditions, where both DNA and protein precipitate. After lysis, a neutralization buffer is added. This buffer allows plasmid to renature and stay in the solution, while chromosomal DNA and proteins still precipitates. A centrifugation step separates the precipitate and the supernatant containing the plasmids. The supernatant was transferred to a spin-column with a collection tube. This spin column is subsequently centrifuged and washed, before transferred to a new tube. Finally, the plasmid DNA is eluted. Elution can be done by the use of elution buffer, or sterile water at neutral or higher pH.

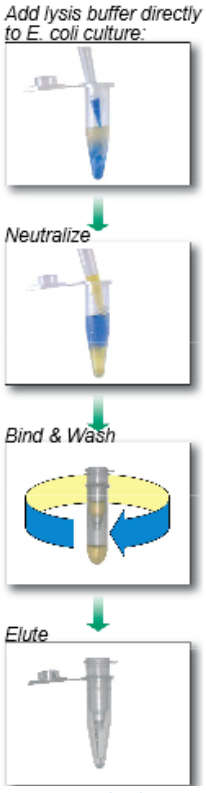



**Figure 14. The basic steps in the Zoppy™ plasmid miniprep kit procedure. Figure is collected from manufacturer's protocol.**

### 3.2.11. Western blot

#### *Measuring protein concentration using Direct Detect™*

Before starting Western blot (WB), protein concentrations were measured using Direct Detect™ from Millipore. Direct Detect™ is an infrared-based quantitation device that measures protein concentration from a slide containing dried 2 µl cell lysate.

#### *Sodium dodecyl sulfate polyacrylamide gel electrophoresis*

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a method used to separate proteins based on molecular mass. Sodium dodecyl sulfate is a detergent that binds to hydrophobic regions, causing them to unfold and release them from associations with other molecules [20].

Calculations and adjustments were performed to ensure that the same protein concentration was used in every well of the SDS-PAGE. Five µl 4x NuPAGE LDS Sample Buffer and 2.5 µl NuPAGE sample reducing agent was added to 20 µl of each sample, bringing the total volume up to 27.5 µl. This was then vortexed and placed in a MBS 0.2G Thermal hybaid for 10 minutes at 99°C for denaturation.

XCell SureLock™ Mini-Cell (Invitrogen) was used for protein separation by SDS-PAGE and blotting. Manufacturer's protocol was followed, with the following exceptions:

- Precast NuPAGE® Novex Bis-Tris 4-12% gels were used
- Appropriate running buffer for this gel (see above) was used (NuPAGE®MES SDS Running Buffer)
- The SDS-PAGE gels were run for 1.5 hours at 200V.

NuPAGE® Antioxidant was added to the inner chamber to avoid re-oxidation of the reduced proteins during migrations, since the reducing agent does not co-migrate with the samples. The gels were run in 1xNuPAGE running buffer (Invitrogen).

### *Blotting*

The proteins were transferred from the gel over to polyvinylidene difluoride (PVDF) membranes, where they are immobilized. The procedure is called "blotting" and transfers the proteins by the use of electric current. The proteins of interest can then be detected by the use of antibodies. Antibodies used are listed in Table 13.

**Table 13 an overview of antibodies used during Western Blot.**

| | Type | concentration | Dilution | Target | Company |
|---|---|---|---|---|---|
| **Primary antibody** | **Rabbit polyclonal IgG** | **200 µg/ml** | **1:200** | **BRCA1 (D-20)** | **Santa Cruz Biotechnology** |
| | | **1 mg/ml** | **1:1000-7000** | **HA-tag** | **Abcam** |
| **Secondary antibody** | **IgG Chicken-anti rabbit - HRP** | **400 µg/ml** | **1:40.000** | **Rabbit antibodies** | **Santa Cruz Biotechnology** |

Blotting procedure was done according to manufacturer's protocol. In brief, the membrane was hydrated by treating it with methanol (30 sec) and MilliQ water (10 sec), and then placed in transfer buffer (appendix B) until used. The gel was then placed between blotting pads and membrane in the following order: 2 blotting pads + filter paper (originally surrounding membrane) + gel (protein side towards membrane) + membrane + filter paper + 3 blotting pads. This was then placed in the XCell SureLock™ Mini-Cell, the inner chamber was filled with transfer buffer and outer chamber with cold water for cooling. The blotting conditions were 1.5 hour at 25V.

After blotting:

1. The membranes were washed for 10 minutes in PBS (appendix B) shaking.
2. In order to block the remaining hydrophobic binding site and avoid non-specific binding of the antibodies, the membranes were transferred to 15 ml blocking buffer (appendix B), and left on the shaker for minimum 1 hour at RT.
3. Four ml blocking buffer and appropriate amount of primary antibody was mixed in a 15 ml tube (see Table 13), before the membranes were transferred. The tubes were put on a rotor (Stuart Rotator SB3, Bibby Scientific Limited) at 4°C overnight.
4. The next day the membranes were washed 3 x 5 minutes in 15 ml PBST (0,1% Tween)
5. Thirty ml PBST was mixed with secondary antibody (see Table 13), producing a 1:40 000 dilution (0.75 µl secondary antibody). Secondary antibodies binding primary antibodies are conjugated to Horseradish peroxidase.
6. Membranes were incubated in the mixture containing the secondary antibodies for 1 hour at RT (shaking).

7. Subsequently membranes were washed 4 x 5 min in 15 ml PBST
8. The membranes were put in a plastic bag and treated with Super Signal® West Dura Extended Duration Substrate (500 µl stable peroxidase solution + 500 µl Luminol/Enhancer solution).
9. After addition of substrate, the membranes were incubated for 5 minutes in the dark.
10. The excess fluid was dried off and the plastic bag was sealed. The membranes were kept in the dark for 10 minutes of incubation before exposure and visualization of the proteins.
11. The complex was visualized on BioDoc-IT from UVP.

## 3.3 Mammalian cell culture techniques

### 3.3.1. General information

In this thesis A375 and HEK-293 cells, which are a malignant melanoma and human embryonic kidney cells respectively, were used. A375 is a cell line derived from a 54 year old female with malignant melanoma carrying the common mutation V600E in B-RAF, while HEK-293 is embryonic kidney cells transformed by adenovirus.

All cultures were grown in DMEM (Dulbecco's Modified Eagle Medium) with 10% FBS (fetal bovine serum) and the antibiotics penicillin and streptomycin (mixture in appendix B). Cell cultures were grown in an incubator set at 37°C with 5 % $CO_2$ (Hera cell 150i, from Thermo Scientific). All reagents where preheated to 37°C before use.

### 3.3.2. Sub-culturing cells

Adherent cells grow in monolayers. When these cells become confluent, they will either die, overgrow each other or both. Dead cells poses a threat for the other cells by releasing harmful cell components into the growth media, and overgrown cells do not have an easy access to nutrients. To overcome these problems, cells can be sub-cultured before they become confluent. In brief, the former growth medium is aspirated, and the cells are washed twice with PBS. PBS is used to remove dead cells and medium. Trypsin was then used to cleave extra cellular proteins linking the cells to the surface of the culture flask. After the enzymatic treatment with trypsin, the detached cells were resuspended in growth medium to inhibit trypsin activity. A small amount of the resuspended cells were transferred to a larger volume of growth media. The A375 cells only needed a short trypsin treatment to detach. Table 14 contains the volumes of reagents needed for sub-culturing cells in medium and large flasks. The growth medium was mixed in advance. Commonly, cells were subcultured and diluted 1:10 or 1:20 for A375, and 1:5 or 1:10 for HEK293, depending on how confluent the cultures were before trypsin treatment.

**Table 14 Volumes of reagents used for sub-culturing cells in medium and large culture flasks.**

| Reagents | Medium culture flask | Large culture flask |
|---|---|---|
| **Growth medium** | ~15 ml | ~30 ml |
| **FBS** | 10 % | 10 % |
| **PBS** | 10 ml | 20 ml |
| **Trypsin** | 1,5 ml | 1,5 ml |
| **Growth medium for resuspension** | 10 ml | 10-15 ml |

### 3.3.3. Seeding cells

Cells were seeded for transfection in 6 well plates. Depending on further experiments, the amount of cells seeded out in each well might differ. Haemocytometric counting was used to determine amount of cells. In brief, cells were treated with trypsin until cells detach. Subsequently, trypsin was inhibited by adding medium containing 10 % fetal bovine serum. Resuspended cells were counted with the help of a Bürkner counting chamber and light microscope. The average count per square times $10^4$ gives an estimate of the number of cells per ml. The volume correction factor $10^4$, is used since each square is 1 x 1 mm and the depth is 0.1 mm (=$10^{-4}$ $cm^3$ or $10^{-4}$ ml).

*# Cells per ml = (summed total of cells/ 9 squares) x $10^4$*

Three hundred thousand HEK293 cells, or 200,000 A375 cells were seeded per well in a 6-well plate. Numbers are based on growth rate.

### 3.3.4. Transient transfection

Transient transfection is a method used to introduce foreign DNA into the cell, without integrating the DNA in the cellular genome. Cells, A375 and HEK-293, were seeded out in a 6-well plate. For the plate with HEK-293 300,000 cells per well was used, since these cells are slow growers compared to A375 cells. For A375 only 200,000 cells were used.

When cells have become 70-80 % confluent, they are transfected using Fugene (Promega) and Opti-MEM® (Invitrogen). Two µg plasmid DNA was mixed with 6 µl Fugene, then 100 µl Opti-MEM was added and mixed. This solution is then left for 15 minutes at RT. In the meantime, cells are washed twice with 2 ml PBS per well and subsequently added 2 ml growth medium. After 15 minutes, the transfection-mix was added dropwise to its respective wells. An overview of the transfection scheme of a 6-well plate is illustrated in Figure 15.

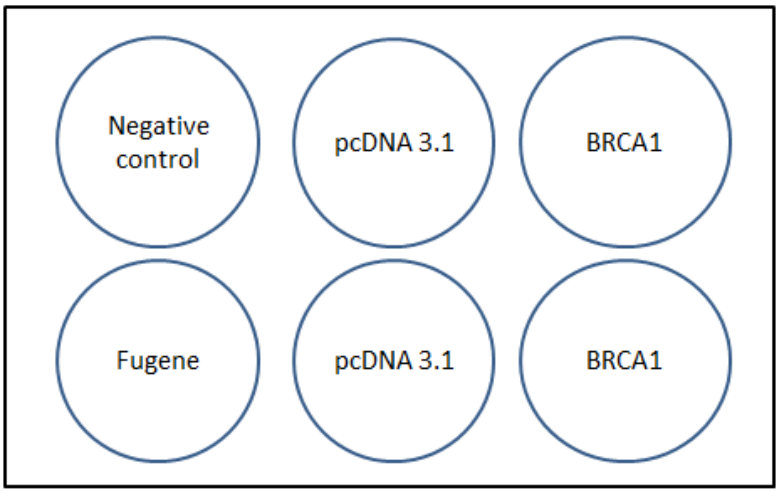



**Figure 15 A simple overview of the transfection of cells in the 6-well plate. Negative control supplied exclusively with growth medium; the well labeled Fugene is provided Fugene and Opti-MEM, but no plasmid DNA; pcDNA 3.1 is a well which is transfected with 2 µg pcDNA 3.1 DNA; whereas wells labeled with BRCA1 are transfected with 2 µg pcDNA3.1 HA-BRCA1 Ser1387Ala plasmid DNA.**

### *Cell protein collection*

For harvesting the cells after transfection, M-PER® Mammalian Protein Extraction Reagent (Pierce Biotechnology) was used. M-PER had to be prepared by adding a tablet of Complete Mini EDTA-free Protease Inhibitor Cocktail Tablets (Roche). One tablet was dissolved in 7 ml M-PER and left at RT for 15 minutes. Meanwhile, wells were washed twice with 2 ml PBS per well. Two hundred µl M-PER with proteinase inhibitor was added dropwise to each well. Cell lysates were gathered with a scraper and transferred into eppendorf tubes that were immediately put on ice.

The cell lysates were centrifuged at 4°C for 10 minutes at 13,000 rpm (Microfuge™ 11, Beckman Coulter) to remove cell debris. The supernatant was transferred to a new tube and put on ice until protein concentration was measured.

## 3.4. Bioinfomatic tools

### 3.4.3. Sequencher 4.10.1

To view the sequencing results, a bioinformatic tool is required. In this thesis we used Sequencher version 4.10.1. Sequencher enables one to evaluate the sequence quality, but it also provides the possibility to manually overwrite the base-calls if bases were faulty called. This software was used to evaluate all the sequences generated during this study.

The Sequencher program automatically compares forward and reverse complement orientations.

Assembly parameters used:

Assembly algorithm: dirty data

- This is a more rigorous alignment algorithm, but it allows untrimmed sequences to be aligned.

Minimum match percentage: 90

- This was default setting. Ensures that sequences have at least a match of 90 % before put into the same contig.

Minimum overlap: 18

- Minimum overlap of bases needed for sequence to be put in the same contig. Shorter overlap leads to placement in a different contigs. Default setting here is 20, but we put it lower so no primers were excluded based on short sequences.

Both "Use ReAligner" and "Prefer 3'Gap Placement" is checked.

- These are default settings. "Use ReAligner" enables user to do a realignment to remove manually produced gaps during handling of sequences.

### 3.4.2. Alamut

Alamut v.2.2 rev.0 is a program which is linked to several prediction programs, databases and related literature. Alamut was used to predict possible effects (splicing and others) of the variants, and to see what has been reported before.

Alamut v.2.2. rev.0 contains links to different relevant databases. The two most relevant for our study are dbSNP and HGMDp. dbSNP is a database containing the majority of reported variants, including both pathogenic variants and polymorphisms. HGMDp is a database of variants that have been published and have been associated with certain diseases.

Alamut is linked to several prediction programs. These programs are divided into predictions for the amino acid substitutions effect on the protein, and for nucleotide substitutions effect on splicing. There are 4 different links for the prediction of sequence variants effect on the protein: Sorting Intolerant From Tolerant (SIFT) [56, 57]; Align Grantham Variation Grantham Deviation (AlignGVGD) [58, 59]; MutationTaster [60]; PolyPhen2 [61]. The latter is not directly linked in Alamut, but Alamut automatically writes the query when using the provided link. SIFT uses multiple alignment information to predict if the substitution is tolerated or deleterious based on conservation throughout evolution [57]. AlignGVGD combines biophysical characterization of amino acids and protein multiple sequence alignments to predict if a missense substitution is deleterious or neutral [59]. MutationTaster is able to predict the disease causing potential of DNA sequence alteration, evaluating evolutionary concervation, splice-site changes, loss of protein features and changes that might affect amount of RNA. MutationTaster chooses between 3 different prediction models based on the nature of the alteration, and in the case of the link with Alamut: alterations concerning a single aa change [60]. Polyphen2 is divided in HumDiv and HumVar, the difference being the corresponding training sets. HumDiv is trained on alleles causing human Mendelian diseases, while HumVar is trained on all disease causing alleles and non-synonymous SNP [61].

For nucleotide substitutions effect on splicing, Alamut has included five splicing prediction programs, two ESE prediction programs and a prediction of branch-point. The splicing prediction programs includes: Splice Site Prediction by Neural Network (NNsplice), Splice Site Finder-like (SSF-like), MaxEntScan (MES), Human Splicing Finder (HSF) and Gene Splicer. The NNsplice, SSF-like and MES use knowledge about base composition of the splice site consensus. All 3 generate a scoring matrix for donor and acceptor site, and SSF-like also includes a scoring matrix for branch-point. Gene Splicer combines several splice site detection techniques, among which Markov models. HSF is based on position weight matrices, but seems to readily predict splice sites without confirmation from the other programs. Human Splicing Finder is therefore not put much weight on during evaluation of splice-signals in this current study. The higher the scores from these programs the more certain the predictions [62].

As mentioned above, Alamut also includes two ESE predictions programs: Relative Enhancer and Silencer Classification by Unanimous Enrichment (RESCUE-ESE) and ESE Finder. However, these predictions, together with branch point prediction reveal a huge number of putative sites, and are therefore not put much weight on during evaluation of splice-signals in this current study [62].

### 3.4.3. Designing primers using Primer3

Primer3 is a web-based program that provides a tool for selecting primers depending on given parameters. This program is found at http://frodo.wi.mit.edu/ (v. 0.4.0), and all primers designed in this study was designed with this program. To avoid primers with multiple binding sites, human mispriming library was used. In addition, a proper identification of primers was given and primer melting temperature (Tm) was set to 59˚C, 60˚C and 61˚C, for minimum, optimum and maximum temperatures, respectively. The sequence to pick primers from was provided to the program in a 5'-3' direction and sequence that primers should not bind to was marked with square brackets at both sides. The rest of the options were left at default.

If possible, selected primers were more than 70 nucleotides from the mutations in *BRCA1* and *BRCA2* to ensure good reads after sequencing.

# 4. Results

## 4.1. *PALB2*

One of the purposes of this study was to investigate if mutations in *PALB2* are a common cause for HBOC in the northern-Norwegian population. To screen for *PALB2* mutations, an efficient method for sequencing using M13 universal primers was established.

### 4.1.1. Test of M13 primer based sequencing

M13 primer based sequencing is a more efficient sequencing method, since it allows us to sequence several distinct exons using the same M13 primers. Necessary for this sequencing method is that the template-specific primers used for amplification have flanking M13 sequences. This means that the amplicons generated start and end with a M13 specific sequence. One M13 specific sequence for forward reads and one for reverse reads. The M13 based sequencing method also reduces the risk of pipetting errors, compared to having several individual primers used for sequencing.

We started out using a DNA sample from a normal control, and sequenced exon 1-13 of *PALB2* in order to establish this sequencing method. Later we focused on only two exons/ one *PALB2* specific primer set (exon 2 and 3) to reduce sequencing cost during development of this method. Two different PCR-programs were used: BRCA1_2_PCR.HYB and BRCA1_2_PCR_20s.HYB (listed in appendix A). As a control we used the *PALB2* specific primers in the sequencing reaction.

#### *M13 complement primers purchased from Sigma-Aldrich*

The first primers tested were "M13 complement" primers purchased from Sigma-Aldrich (Table 9). The initial primers had been purified by desalting. . PCR products amplified by RedTaq and sequenced with these M13 primers resulted in indecipherable sequences. There were only 3 sequences with a quality score above 50 %. The quality score provided by the sequencer software is an indication of how sure the base calls are throughout the sequence length. However, the PCR-products sequenced with exon-specific primers (the same as in the amplification step) gave readable sequences.

#### *Test of HotStarTaq plus DNA polymerase*

The former results led us to test a different polymerase for the amplification step, in order to exclude if some of the components of the RedTaq-mix might interfere with the M13 based sequencing. As an alternative to RedTaq, HotStarTaq plus DNA polymerase was used to generate PCR products. However, also the latter generated PCR-products that could not be sequenced with M13 primers. A representative of the unreadable sequences is illustrated in Figure 16.

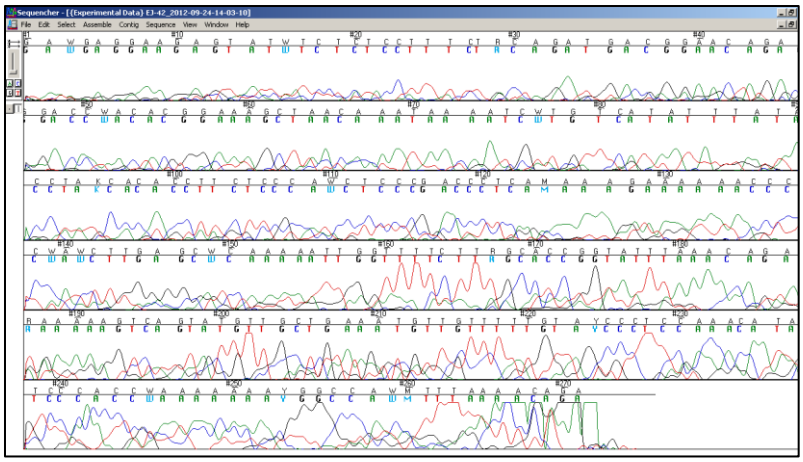


**Figure 16 An electropherogram of the sequences generated from a PCR products amplified with HotStarTaq plus DNA polymerase instead of RedTaq and sequenced with the desalted M13 primers purchased from Sigma-Aldrich.**

### *Test of template concentration and primers from kit*

To investigate whether template concentrations in the sequencing reaction might improve the results, we used 1 µl, 2 µl and 4 µl PCR-product as template (Table 10). Simultaneously, different M13 primers were tested. The M13 sequencing primers used were provided with a BigDye direct cycle sequencing kit (Life Technologies). M13 primer concentrations were not provided with the kit, but the reaction mix proposed by the manufacturer was 1 µl primers in a total volume of 13 µl. Accordingly, we used 1.5 µl primers in 20 µl total sequencing reaction mixture.

All 3 different sequencing reactions, with different amounts of template were able to successfully sequence the PCR products used.

In a subsequent study to investigate if the polymerase had any influence we retested PCR products produced by RedTaq. Results were equal to the sequencing of PCR-product produced with HotStarTaq (Figure 17).

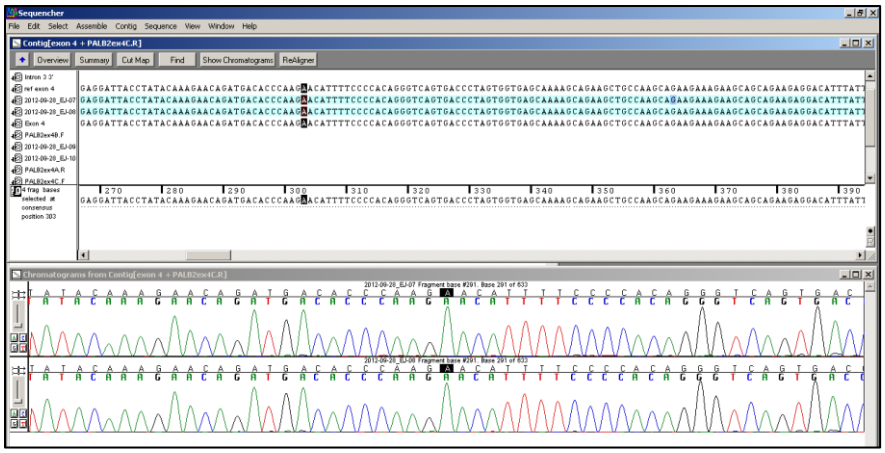


**Figure 17 Sequences produced with M13 primers provided in a BigDye Direct Cycle Sequencing Kit (Applied Biosystems). This screenshot represents sequences generated from PCR products amplified by RedTaq polymerase and analyzed using Sequencher 4.10.1 Software. In the top part of the figure sequences are shown in light green and represent sequences of good quality. The lower part of the figure is the electropherogram of the same sequence.**

### *M13 primers purchased from Life Technology – primer concentration*

As a consequence of the previous results, M13 primers were ordered from Life Technologies. Since the concentration of the M13 primers was not provided by the manufacturer in the BigDye direct cycle sequencing kit previously used, different M13 primer concentrations were used in the subsequent optimization of the sequencing reactions. According to manufacturer's recommendations, these primers should be diluted to 20 µM (pmol/µl) prior to use. We tested different final concentrations of M13 pimers in the sequencing reaction ranging from 25nM to 1µM. No obvious difference in sequencing quality was observed. In order to make M13 based sequencing even more cost efficient, the total volume of our sequencing reaction was reduced from 20 to 10 µl without any obvious impact on our sequencing results (Figure 18).

The M13 primer based sequencing method development was now considered satisfactory, and this method was subsequently used for *PALB2* sequencing.

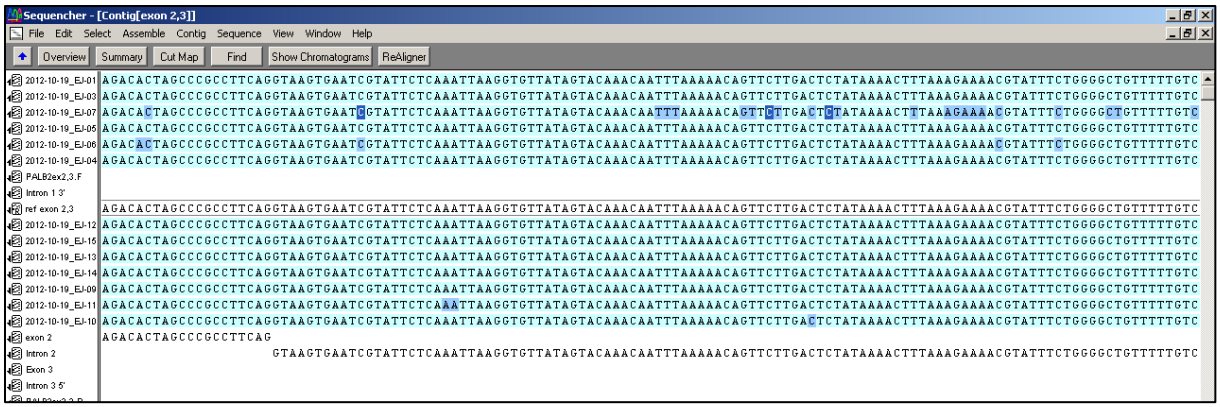



**Figure 18 Sequence generated from a 10 µl sequencing reaction mix (BigDye v.3.0) with 25nM final M13 primer concentration. M13 primers are purified by desalting and subsequent gel electrophoresis and purchased from Life Technologies. RedTaq was used for DNA amplification. Confident base calls are displayed in light green using Sequencher software.**

### 4.1.2. Screening of *PALB2* for cancer causing mutations

In total, 43 patients were included in this part of the study (Table 19 in appendix A). Thirty five patients were diagnosed with familial breast cancer, 7 were diagnosed with familial ovarian cancer and one patient had both breast and ovarian cancer. We investigated all 13 exons in *PALB2* for mutations, including intron/exon boundaries.

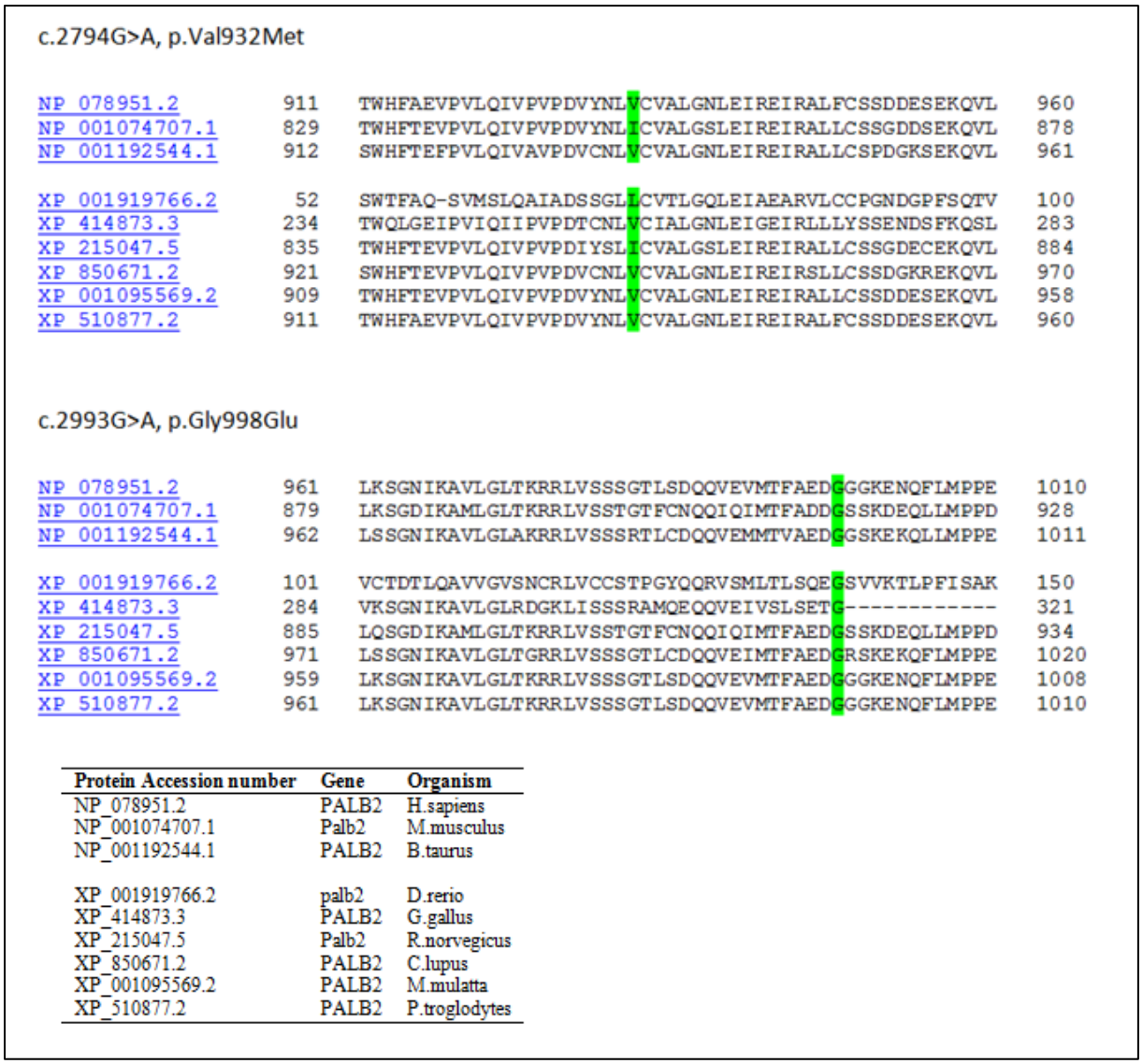
When screening *PALB2* for mutations 14 different variants were identified. Six of these were located in 5 different exons, 1 was located in the 5'UTR, while the remaining 7 were located in 6 different introns. A summary of the results is illustrated in Table 15**Feil! Fant ikke referansekilden.**. The significance of the identified variants was assessed with Alamut v.2.2rev.0 software. Predictions from SIFT, AlignGVGD, Polyphen2 and MutationTaster, together with access number for dbSNP and HGMDp are listed in Table 15. Variants are classified into 5 classes:

1: benign
2: likely benign
3: variant of unknown clinical significance (VUS)
4: probably pathogenic
5: pathogenic

Five of the *PALB2* mutations were classified as VUS including; two exonic variants, c.2794G>A and c.2993G>A; two intronic variants, c.2748+92T>G and c.2749-18C>T; and one mutation located in the 5'UTR c.-5G>C. Of the 2 exonic variants classified as VUS, both were classified as class C0 by AlignGVGD, but deleterious by SIFT and probably damaging by PolyPhen2. MutationTaster

classified c.2794G>A as a polymorphism and c.2993G>A as disease causing. Little information was available for the intronic variant c.2748+92T>G and the 5'UTR c.-5G>C. Only the c.2749-18C>T of the two has been previously reported in the literature (Table 15).

One of the variants detected during screening was the ancestral allele; c.3114-51A. All samples screened for mutations in *PALB2* were homozygous for c.3114-51T>A and this variant is subsequently classified as class 1. Three other variants, c.1676A>G, c.2996+264T>C and c.3201+101AG, are also classified as class 1 (benign), mostly based on minor allele frequency (MAF) (Table 15). The remaining variants, c.1010T>C, c.2014G>C, c.3300T>G, c.212-58A>C, c.2586+58C>T, are all classified as class 2: likely benign (Table 15). Another thing we can see from Table 15 is that some of the variants are more represented than others, in our group of 43 patients. In addition, none of the exonic variants are found in a homozygous state.




| Protein Accession number | Gene | Organism |
|---|---|---|
| NP_078951.2 | PALB2 | H.sapiens |
| NP_001074707.1 | Palb2 | M.musculus |
| NP_001192544.1 | PALB2 | B.taurus |
| XP_001919766.2 | palb2 | D.rerio |
| XP_414873.3 | PALB2 | G.gallus |
| XP_215047.5 | Palb2 | R.norvegicus |
| XP_850671.2 | PALB2 | C.lupus |
| XP_001095569.2 | PALB2 | M.mulatta |
| XP_510877.2 | PALB2 | P.troglodytes |

**Figure 19 Amino acid alignment from HomoloGene at [www.ncbi.nlm.gov/](www.ncbi.nlm.gov/) of the two exonic variants classified as VUS. Only three sequences were available from HomoloGene. The rest are predicted sequences (XP). Species are listed in table under the alignments. Amino acids (represented with single letter code) marked in a grey bar correspond to the position of amino acids changed by the missense substitution for the two *PALB2* variants: c.2794G>A (Val932Met) and c.2993G>A (Gly998Glu).**

**Table 15. Overview of the mutations found in *PALB2*. The Single Nucleotide Polymorphism database (dbSNP) is an archive for simple genetic polymorphisms, and the frequencies by which they occur (Minor Allele Frequencies (MAF)). The Human Genome Mutation Database (HGMDp) comprises every previously described disease associated mutations. Sorting Intolerant From Tolerant (SIFT) and Align Grantham Variation, Grantham Deviation (AlignGVGD) are both prediction programs based on the evolutionary conservation at their location within the protein, while AlignGVGD also takes into consideration amino acid characteristics. Polyphen2 is divided in HumDiv and HumVar, difference being the training sets. MutationTaster prediction is based on evolutionary conservation, splice-site changes, loss of protein features and changes that might affect amount of RNA. Classification is based on both the prediction programs and mutation database and literature searches. Nomenclature is according to HGVS standards, c.1 corresponds to the first nucleotide in ATG.**

| Mutation | Location | Type | Protein | # of patients | dbSNP | MAF | HGMDp | SIFT | Align GVGD | PolyPhen2 HumDiv | PolyPhen2 HumVar | Mutation Taster | Homozygous patients | Heterozygous patients | Class | Ref. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Exonic:** | | | | | | | | | | | | | | | | |
| **c.1010T>C** | exon 4 | Missense | p.Leu337Ser | 2 | rs45494092 | 0.012 | CM112098 | Deleterious (0.00) | C0 | Benign (0.291) | Benign (0.117) | Polymorphism (0.952) | | 5, 19 | 2 | [15, 63-66] |
| **c.1676A>G** | exon 4 | Missense | p.Gln559Arg | 8 | rs152451 | 0.146 | - | Tolerated (1.00) | C0 | Benign (0.000) | Benign (0.000) | Polymorphism (1.00) | | 5, 9, 19, 23, 25, 37, 38, 39 | 1 | [15, 63-68] |
| **c.2014G>C** | exon 5 | Missense | p.Glu672Gln | 1 | rs45532440 | 0.014 | CM112099 | Deleterious (0.04) | C0 | Benign (0.225) | Benign (0.047) | Polymorphism (0.98) | | 23 | 2 | [15, 64, 65, 67, 69] |
| **c.2794G>A** | exon 8 | Missense | p.Val932Met | 1 | rs45624036 | 0.001 | CM112100 | Deleterious (0.01) | C0 | Probably damaging (1.000) | Probably damaging (0.993) | Polymorphism (0.783) | | 16 | 3 | [15, 63, 66, 68] |
| **c.2993G>A** | exon 9 | Missense | p.Gly998Glu | 2 | rs45551636 | 0.011 | CM098533 | Deleterious (0.00) | C0 | Probably damaging (1.000) | Probably damaging (0.999) | Disease causing (0.953) | | 23, 39 | 3 | [15, 64, 66, 67] |
| **c.3300T>G** | exon 12 | Synonymous | p.= (p.Thr1100Thr) | 2 | rs45516100 | 0.018 | - | - | - | - | - | - | | 23, 39 | 2 | [15, 63-65, 67-69] |
| **Intronic:** | | | | | | | | | | | | | | | | |
| **c.212-58A>C** | intron 3 | | p.? | 2 | rs80291632 | 0.014 | - | | | | | | 23(?) | 23(?), 39 | 2 | [64, 65, 67, 69] |
| **c.2586+58C>T** | intron 6 | | p.? | 17 | rs249954 | 0.335 | - | | | | | | 9, 25, 39 | 3, 4, 12, 18, 19, 23, 24, 27, 30, 33, 37, 38, 41, 43 | 2 | [64-66, 69, 70] |
| **c.2748+92T>G** | intron 7 | | p.? | 2 | - | - | - | | | | | | | 7, 9 | 3 | |
| **c.2749-18C>T** | intron 7 | | p.? | 1 | rs182194007 | - | - | | | | | | | 16 | 3 | [64] |
| **c.2996+264T>C** | intron 9 | | p.? | 18 | rs420259 | 0.345 | - | | | | | | 9, 25, 39 | 3, 4, 12, 18, 19, 23, 24, 27, 30, 33, 37, 38, 40, 41, 43 | 1 | [70] [64] |
| **c.3114-51T>A** | intron 10 | | p.? | 43 (all) | rs249936 | 0.000* | - | | | | | | all patients | | 1 | [64] |
| **c.3201+101A>G** | intron 11 | | p.? | 5 | rs249935 | 0.146 | - | | | | | | | 4, 9, 19, 23, 25 | 1 | [64] |
| **UTR:** | | | | | | | | | | | | | | | | |
| **c.-5G>C** | 5'UTR | | p.? | 1 | - | - | - | | | | | | | 28 | 3 | |

* Ancestral allele

## 4.2. *BRCA1* and *BRCA2*

In the second part of this study we investigated variants of unknown clinical significance identified in *BRCA1* and *BRCA2*. These variants were tested for alternatively spliced transcripts using cDNA from whole blood. In addition, we investigated if both alleles were transcribed in blood.

### 4.2.1. RNA purification from whole blood

Nineteen patient samples with VUS in *BRCA1* and 18 patient samples with VUS in *BRCA2* were investigated, corresponding to 20 *BRCA1* variants and 15 *BRCA2* variants (see Table 1 and Table 2 in methods). RNA from patient samples was purified from blood collected in PAXgene tubes, while the normal control (sample number 120483) was isolated from whole blood stored in a tempus tube.

After RNA isolation, the RNA concentrations were measured using NanoDrop. Concentrations ranged from 30-212 ng/µl (median 70 ng/µl) with one exception of 8ng/µl (sample 73). For the latter a modified reaction mix was used during cDNA synthesis (Table 4 in methods).

After cDNA synthesis, all samples were tested using the APRT-test. A distinct band is present at 218bp, corresponding to amplification of cDNA using the APRT-primers. An example of gel electrophoresis of APRT can be seen in Figure 20. A weaker, higher molecular mass band is also present, corresponding to amplification of gDNA with the same primers. The difference in size equals the length of introns incorporated in the amplification of the gDNA. Gel images from patient samples can be seen in Figure 33 and Figure 34 in Appendix C.

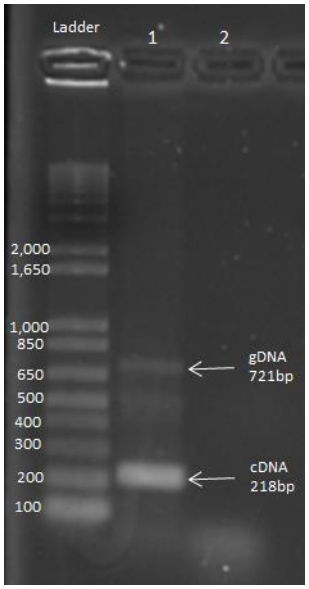



**Figure 20. A visualization of the PCR-products from the APRT–test. The ladder is 1 Kb plus DNA molecular marker (Invitrogen). (1) Normal control (sample 120483). (2) No template control.**

### 4.2.2. Results from transcript analysis

The *BRCA1* and *BRCA2* variants of unknown clinical significance were tested for aberrant splicing and tested to see if both alleles were transcribed. Since these mutations were identified in genomic DNA and patients were heterozygous for them, it is possible to use them as an indicator to see if both alleles are transcribed. This is only possible for exonic variants, unless the part of the intron containing the variant is included due to aberrant splicing.

For examination of cDNA sequences, two projects (one for *BRCA1* and one for *BRCA2*) were made in Sequencher. The project contained the cDNA sequence, 3' and 5' UTR, and primers used for amplification. This could then be imported when sequencing results became available. This simplified the analysis of sequences for possible alternative transcripts.

In order to investigate if mutations undetectable by sequencing of genomic *BRCA1* and *BRCA2* were present, cDNA of all patients carrying a variant of unknown significance were examined for the presence of both alleles in blood. There were no indications of gDNA-undetectable mutations (see Table 16 and Table 17), but for samples carrying intronic variants it was not possible to conclude if both alleles were transcribed. Also, for three samples/two variants we did not manage to amplify enough of the target-sequence for subsequent sequencing (*BRCA2* c.3568C>T and *BRCA2* c.4068G>A).

**Table 16 Mutation variants located in *BRCA1* tested for alternative splicing and transcription of both alleles. Gel images can be found in figure 35-41 in appendix C. Classification is based on results from prediction programs and mutation databases and literature searches. Splicing is also taken into consideration for variants that seem to affect splicing. Alternative splicing which was detected both in samples and in the normal control was not included in this table. lmw = low molecular weight compared to main/expected band, hmw = high molecular weight compared to main/expected band. Class. = classification where 1 is benign, 2 likely benign, 3 variant of unknown clinical significance, 4 probably pathogenic and 5 pathogenic. Nomenclature is according to HGVS standards, c.1 corresponds to the first nucleotide in ATG. Reference sequence: U14680.1**

| Mutation | Localization | Protein | Sample number | PCR-product visualized on agarose gel | Splicing effect seen from sequencing | Both alleles transcribed | dbSNP | HGMDp | Class | References |
|---|---|---|---|---|---|---|---|---|---|---|
| **c.-33-29delAAAAA** | Intron 1 | - | 44 | same as control, but both have a clear hmw band (~290bp) | None* | - | - | - | 3 | [71] |
| **c.1A>G** | Exon 2 | p.Met1? | 45 | same as control, but both have a clear hmw band (~290bp) | None* | Yes | rs80357287 | CM021503 | 5 | [33] |
| **c.130T>A** | Exon 3 | p.Cys44Ser | 46 | same as control | None | Yes | rs80357327 | CM087364 | 5 | [72-78] |
| **c.140G>T** | Exon 5 | p.Cys47Phe | 47 | Same as control, but also a lmw band (~340 bp) | None | Yes | rs80357150 | CM032549 | 4 | [75, 76] |
| **c.213-5T>A** | Intron 5 | - | 48 | Different bond-pattern (Figure 21) | Inclusion of 59 nucleotides of 3' end of intron 4 | Yes | - | - | 4 | - |
| **c.486G>T** | Exon 8 | p.= (p.Val162Val) | 49 | same as control | None** | Yes | - | - | 2 | - |
| **c.548-17G>T** | Intron 8 | - | 50 | same as control | None** | Yes | rs80358014 | - | 1 | [72, 79] |
| **c.734A>T** | Exon 11 | p.Asp245Val | 51 | same as control | None | Yes | rs80357327 | CM087364 | 3 | - |
| **c.1419C>T** | Exon 11 | p.= (p.Asn473Asn) | 52 | same as control | None | Yes | - | - | 2 | - |
| **c.1487G>A** | Exon 11 | p.Arg496His | 53 | same as control | None | Yes | rs28897677 | CM014323 | 1 | [72, 73] |
| **c.2521C>T** | Exon 11 | p.Arg841Trp | 54 | same as control | None | Yes | rs1800709 | CM004236 | 1 | [72, 73, 80, 81] |
| **c.3418A>G** | Exon 11 | p.Ser1140Gly | 55 | same as control | None | Yes | rs2227945 | - | 1 | [72, 73, 76] |
| **c.3708T>G** | Exon 11 | p.Asn1236Lys | 56 | same as control | None | Yes | rs28897687 | CM994631 | 3 | [36, 73] |
| **c.5075A>C** | Exon 18 | p.Asp1692Ala | 57 | Same as control | None | Yes | - | - | 3 | - |
| **c.5096G>A** | Exon 18 | p.Arg1699Gln | 58 | same as control | None | Yes | rs41293459 | CM034007 | 5 | [36, 72, 82] |
| **c.5117G>C** | Exon 18 | p.Gly1706Ala | 54 | same as control | None | Yes | rs80356860 | CM030790 | 1 | [72, 83, 84] |
| **c.5123C>T** | Exon 18 | p.Ala1708Val | 59 | same as control | None | Yes | rs28897696 | CM065004 | 4 | [85, 86] |
| **c.5125G>A** | Exon 18 | p.Gly1709Arg | 60 | same as control | None | Yes | - | - | 3 | - |
| **c.5434C>G** | Exon 23 | p.Pro1812Ala | 61 | 3 distinct bonds, one clearly different from the normal control (Figure 23) | Exclusion of exon 23 | - | RS1800751 | CM032862 | 5 | [87, 88] |
| **c.5513T>G** | Exon 24 | p.Val1838Gly | 62 | same as control | None | Yes | - | - | 3 | - |

* Observed skipping of GTAAAG in exon 1, also found in normal control

** Observed skipping of CAG in exon 8, and transcripts skipping exon 9 and 10. Also found in normal control.

**Table 17. Mutation variants located in *BRCA2* tested for alternative splicing and transcription of both alleles. Gel images can be found in figure 35-41 in appendix C. Classification is based on results of prediction programs, mutation databases and literature searches. Splicing is also taken into consideration for variants that seem to affect splicing. We were not able to produce decipherable sequences from two variants: c.3568C>T and c.4068G>A. Class. = classification where 1 is benign, 2 likely benign, 3 variant of unknown clinical significance, 4 probably pathogenic and 5 pathogenic . Nomenclature is according to HGVS standards, c.1 corresponds to the first nucleotide in ATG. Less distinct bands than primary/expected band are here termed "shadow bands". Reference sequence: NM_000059.1**

| Mutation | Localization | Protein | Sample number | PCR-product visualized on agarose gel | Splicing effect seen from sequencing | Both alleles transcribed | dbSNP | HGMDp | Class. | References |
|---|---|---|---|---|---|---|---|---|---|---|
| **c.40A>G** | Exon 2 | p.Ile14Val | 63 | Same as control | None* | Yes | - | - | 2 | - |
| c.68-7T>A | **Intron 2** | - | **64** | **Weaker main band, stronger second compared to control** | **Increased exon 3 skipping** | - | **rs81002830** | **CS033491** | **3** | **[88, 89]** |
| **c.750G>A** | Exon 9 | p.= (p.Val250Val) | 65, 76 | Same as control | None | Yes | rs143214959 | - | 3 | - |
| **c.2680G>A** | Exon 11 | p.Val894Ile | 66 | Same as control | None | Yes | rs28897715 | - | 1 | [72] |
| **c.3568C>T** | Exon 11 | p.Arg1190Trp | 67 | Weak, but same bands as control | ?** | - | rs80358604 | - | 1 | [72, 90] |
| **c.4068G>A** | Exon 11 | p.= (p.Leu1356Leu) | 68, 55 | Weak, but same bands as control | ?** | - | rs28897724 | - | 3 | [91] |
| **c.4828G>A** | Exon 11 | p.Val1610Met | 69 | Same as control | None | Yes | rs80358705 | - | 2 | [92] |
| **c.5272_5274delAAT** | Exon 11 | p.Asn1758del | 70 | Same as control | None | Yes | - | - | 2 | - |
| **c.6100C>T** | Exon 11 | p.Arg2034Cys | 71 | Same as control | None | Yes | rs1799954 | CM994286 | 1 | [72, 73] |
| **c.6821G>T** | Exon 11 | p.Gly2274Val | 72 | Same as control | None | Yes | RS55712212 | - | 2 | [72, 73] |
| **c.7301A>C** | Exon 14 | p.Lys2434Thr | 73 | Same as control | None | Yes | rs80358954 | - | 3 | - |
| **c.7878G>C** | Exon 17 | p.Trp2626Cys | 74 | Same as control, without shadow band | None | Yes | rs80359013 | CM053139 | 5 | [72, 73] |
| **c.8177A>G** | Exon 18 | p.Tyr2726Cys | 75 | Same as control | None | Yes | rs80359064 | - | 4 | - |
| **c.8323A>G** | Exon 18 | p.Met2775Val | 77 | Same as control, without shadow bands | None | Yes | - | - | 3 | - |
| **c.9116C>T** | Exon 23 | p.Pro3039Leu | 78, 79 | Same as control, but different shadow bands | None | Yes | rs80359167 | - | 3 | [93] |

*Observed skipping of exon 3, also found in normal control
** no decent sequences

After *in silico* evaluation, splicing analysis and test for expression of both alleles, variants were classified into 5 different classes based on results:

1: benign
2: likely benign
3: variant of unknown clinical significance
4: probably pathogenic
5: pathogenic

A lot of variants were still classified as class 3: unknown. This includes six *BRCA1* variants and six *BRCA2* variants (listed in Table 16 and Table 17). Of these, four *BRCA1* and one *BRCA2* mutation have never been reported before.

Thirteen variants were predicted to be benign or likely benign, including seven *BRCA1* and seven *BRCA2* variants. *BRCA1* c.548-17G>T, c1487G>A, c.2521C>T, c.3418A>G and c.5117G>C were all classified as class 1, while c.486G>T and c.1419C>T as class 2. Of the seven mutations in *BRCA2*, c.40A>G, c.4828G>A, c.5272_5274delAAT and c.6821G>T were classified as class 2, while c.2680G>A, c.3568C>T and c.6100C>T as class 1. This is based on literature searches, location outside domains, no alternative splicing or indication of previously undetectable mutations (results from prediction programs can be seen in Table 23 and

Table 24 in appendix C).

For variants classified as either class 4 or 5 in *BRCA1*, one variant (c.1A>G) results in a shorter N-terminal end of the protein by 17 aa, due to the use of an alternative start codon. Two of the mutations are located in the RING-domain; c.130T>A and c.140G>T and c.5123C>T is located in the BRCT domain. Another two (c.213-5T>A and c.5434C>G) produce alternative skipping reported in more detail later on (under the heading 4.2.2.1. BRCA1). For *BRCA2* only two variants were classified either as class 4 or 5, where c.7878G>C is classified as pathogenic based on substitution of amino acid in DSS1 binding protein domain and c.8177A>G is classified as likely pathogenic based on substitution of an amino acid located in the OB domain and high conservation of nucleotide and amino acid at this position (data not shown).

#### *4.2.2.1. BRCA1*

By separating the PCR-product of cDNA samples containing *BRCA1* c.213-5T>A and c.5434C>G variants by gel electrophoresis, an aberrant band pattern was visible (Figure 21 and Figure 23). Subsequent sequencing and examination of these two samples in Sequencher identified alternative

splicing. Mutation c.213-5T>A (Figure 21 and Figure 22 - sample 48) induces an inclusion of 59 bases from the 3'-end of intron 5, which results in a frameshift. Mutation c.5434C>G (Figure 23 and Figure 24 - sample 61) induces exon skipping of exon 23, together with the last base in exon 22. The latter gives a 61 base deletion from the transcript, and therefore also results in a frameshift during translation and a truncated protein.

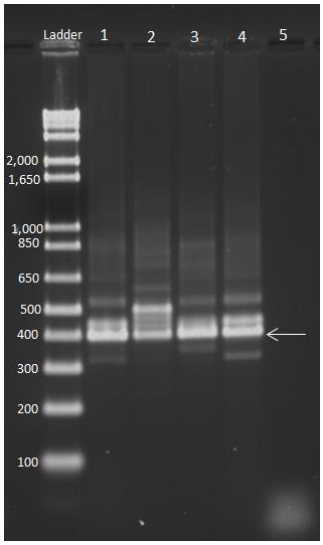



**Figure 21. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen), (1) Normal control, (2) c.213-5T>A (sample 48), (3) c.130T>A (sample 46), (4) c.140G>T (sample 47) and (5) no-template control. The arrow indicates the expected band size from this primer set (BRCA1 130.F and BRCA1 130.R – listed in Table 21 in appendix B).**

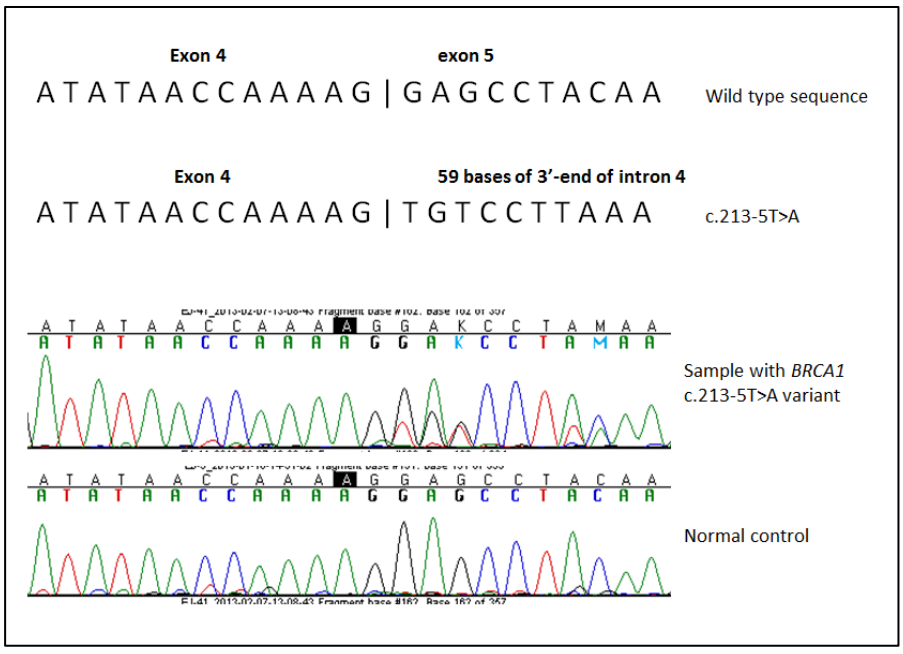



**Figure 22. The mutation c.213-5T>A (Sample 48) resulted in an inclusion of 59 bases from the 3'-end of intron 5 in the investigated cDNA. At the top are bases representing the boundary between exon 4 and 5 in *BRCA1*. The inclusion of**

**intronic bases that c.213-5T>A leads to are represented under the wild type sequence. At the bottom, electropherograms from the** sample **heterozygous for *BRCA1* c.213-5T>A and the normal control are shown.**

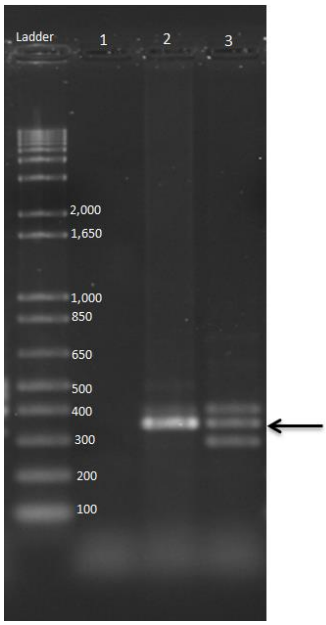



**Figure 23. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen), (1) No-template control, (2) normal control and (3) c.5434C>G (sample 61). The arrow indicates the expected band size from this primer set (BRCA1 5434.F and BRCA1 5434.R – listed in Table 21 in appendix B).**

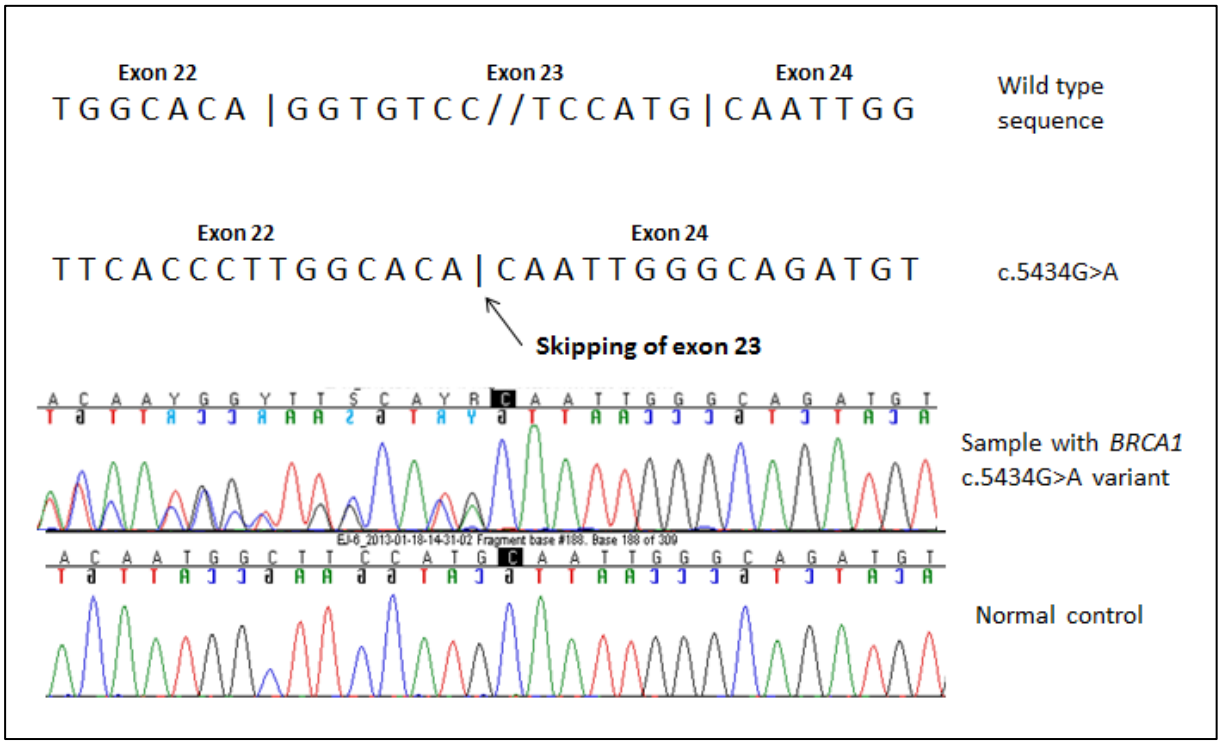



**Figure 24. Mutation c.5434C>G (Sample 61) resulted in skipping of exon 23 in cDNA made from whole blood isolated RNA. Sixty one bases are skipped and accordingly this mutation creates a frame shift during translation. The top represents an overview of bases in a wild type sequence and the skipping of exon 23 that *BRCA1* c.5434G>A induces.**

**At the bottom, electropherograms from the sample heterozygous for *BRCA1* c.5434G>A and the normal control are shown.**

During our investigation of the *BRCA2* variants, only one sample seemed to affect splicing. The results obtained by gel electrophoresis (Figure 25) did not reveal any different band pattern between the control sample and the *BRCA2* variants studied. However, sequencing and examination of sequences in Sequencher, revealed skipping of exon 3 occurring in all samples including c.40A>G, c.68-7T>A, and the normal control. However, c.68-7T>A seems to increase the amount of PCR product which is subjected to skipping of exon 3, resulting in a more intense band of 270bp (illustrated in Figure 25). The B-arrow is indicating the transcripts lacking exon 3, and by comparing this lower molecular weight band (270bp) with the more distinct and expected band (519 bp), you can see a bigger difference than in the normal control.

In addition, we had another good candidate for aberrant splicing, c.9116C>T, based on predictions and gel image as illustrated in Figure 27. However, the additional bands visible on the agarose gel did not result in aberrant sequencing upon sequencing of these PCR-products.

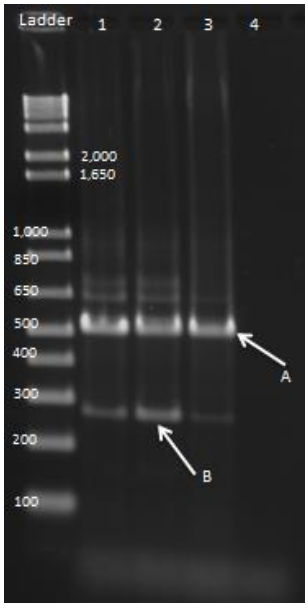



**Figure 25. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen), (1) Normal control, (2) c.68-7T>A (sample 64), (3) c.40A>G (sample 63) and (4) No-template control. The A-arrow indicates the expected band size from this primer set (BRCA2 e1F-ny and BRCA2 e4R-ny – listed in Table 21 in appendix B) and the B-arrow indicates the more intense band, corresponding to increased skipping of exon 3.**

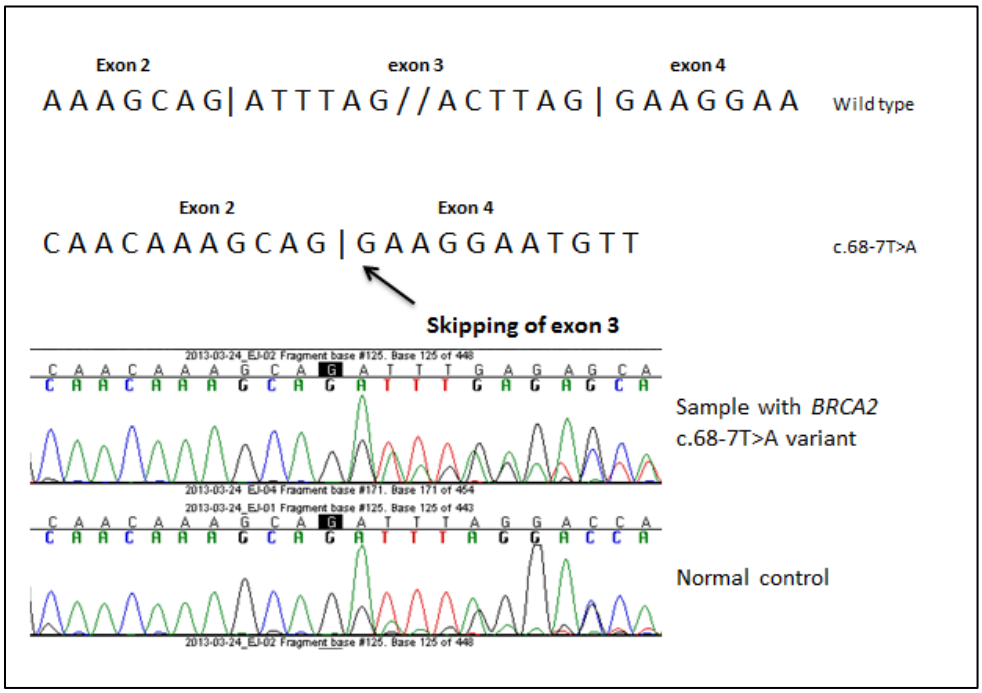



**Figure 26. The mutation c.68-7T>A (Sample 64) resulted in increased skipping of exon 3 in the investigated cDNA. Two hundred and forty nine bases are skipped which do not create a frame shift during translation. The top represents an overview of bases in a wild type sequence and the skipping of exon 3 that *BRCA2* c.68-7T>A seem to increase. At the bottom, electropherograms from the sample heterozygous for *BRCA2* c.68-7T>A and the normal control are shown.**

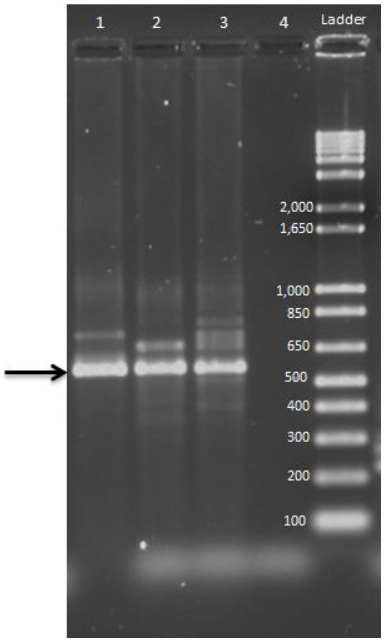



**Figure 27. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen), (1) Normal control, (2) c.9116C>T (sample 78), (3) c.9116C>T (sample 79) and (4) No-template control. The arrow indicates the expected band size from this primer set (BRCA2 9116.F and BRCA2 9116.F – listed in Table 21 in appendix B).**

### 4.3. Normal alternative splicing

Many transcripts have been identified both for *BRCA1* and *BRCA2*. These are the result of alternative splicing and accordingly these transcripts were also found in normal controls. The normal alternative spliced transcripts which were also detected in the present study are listed in Table 18. Three of the detected splice variants were detected in *BRCA1*, one transcript includes the last six bases of exon 1 (ΔGTAAAG ex1), another one excludes the first three bases of exon 8 (ΔCAG ex8) and a third one skips both of exon 9 and 10 (Δ9,10). One alternative spliced variant was also detected in *BRCA2*, this transcript variant exclude the entire exon 3 (Δex3). All these normal splice variants cause in-frame induced changes. As an example of these alternatively spliced variants, an image of the GTAAAG skipping can be seen in Figure 28.

**Table 18. Normal alternatively spliced transcripts.**

| Named here | observed alteration | Number of nucleotides | Gene | Reference |
|---|---|---|---|---|
| **ΔGTAAAG ex1** | skipping of GTAAAG, last nucleotides in exon 1 | 6 | BRCA1 | [94] |
| **ΔCAG ex8** | Skipping of CAG in the start of exon 8 | 3 | BRCA1 | [95] |
| **Δ9,10** | skipping of exon 9 and 10 | 123 | BRCA1 | [96] |
| **Δex3** | skipping of exon 3 | 249 | BRCA2 | [97] |

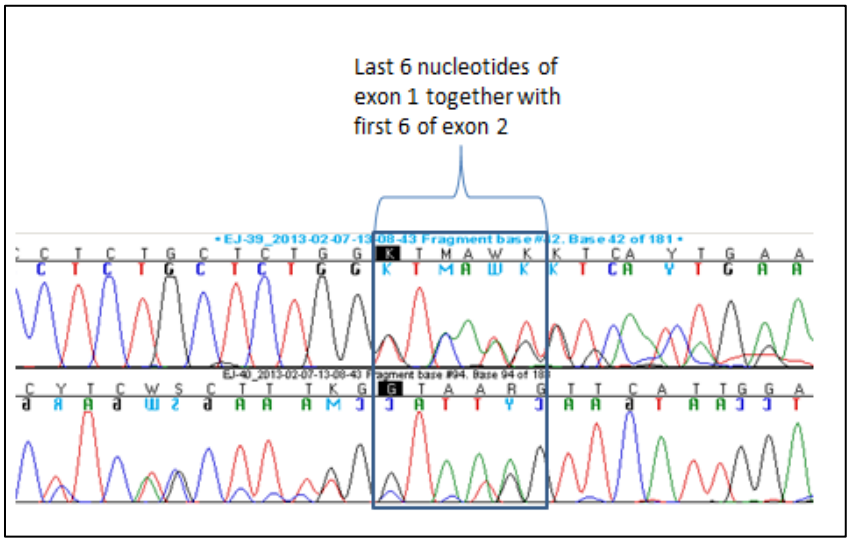



**Figure 28. Exclusion of the last six nucleotides in exon 1 in some of the BRCA1 transcripts. The upper sequence is the forward read and the bottom the reverse over exon 1/exon 2 boundary.**

## 4.3. Effect of VUS on the BRCA1 protein

In order to investigate the influence of the genetic variants in *BRCA1* on protein level, we initiated expression studies of BRCA1 in *in vitro* cell tissue culture.

Mutant pcDNA3 HA-BRCA1 Ser1387Ala plasmid purchased from Addgene was used to generate from pcDNA3 HA-BRCA1 WT by site-directed mutagenesis. After site directed mutagenesis, 14 bacteria clones were checked by PCR. The *BRCA1* insert in the plasmid was checked by PCR producing a 1714bp product. Only a single clone (clone 11, Figure 29) showed an unexpected larger insert. Subsequently, plasmids with the correct PCR product sizes were purified and sequenced. The plasmid was confirmed to now contain the WT *BRCA1*sequence. The pcDNA3-Ha_BRCA1 plasmid was used in the subsequent transfection studies.

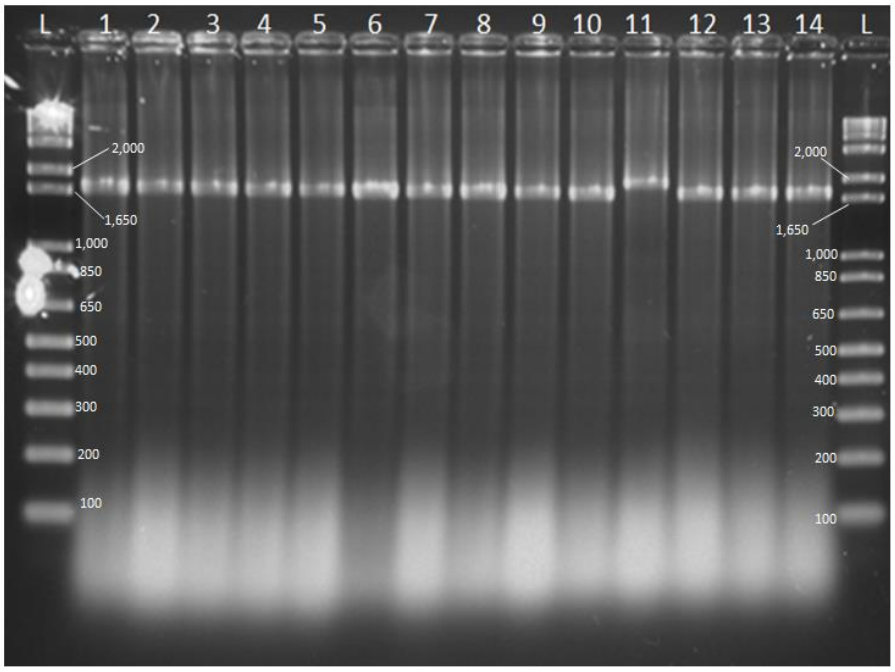



**Figure 29. Visualization of PCR-products from pcDNA3-Ha-BRCA1 WT plasmids. L = 1kb-Plus Ladder (Invitrogen). Products were amplified using primer BRCA1 ex11_fr8.F and BRCA1 5075.R. All lanes except 11 have the right PCR-product size. Sample 11 was removed from further procedures.**

### *Transfection and protein harvesting*

Plasmid pcDNA3 HA-BRCA1 Ser1387Ala was used for transfection of mammalian cell lines. In this study two different cell lines were transfected (A375 and HEK-293). Cell lysates were collected approximately 24 and 48 hours after the transfection (transfection was done in duplicates) for both cell lines. The protein concentration in the cell lysates was determined and is listed in Table 25 in appendix C.

### *Western blot*

During our initial study of the BRCA1 expression in transfected cells, we were unable to detect BRCA1 protein in our cell lysates retrieved from A375 transfected cells (Figure 30). We used BRCA1 (D-20) specific antibodies in this initial western blot analysis. Subsequently, since the transfected plasmid contained an HA-tagged BRCA1 we tried to detect BRCA1 using an HA specific antibody. However, no BRCA1 specific band could be detected compared with the control. Similar results were obtained for western blots performed on HEK-293 cell lysates (data not shown).

Different concentration of primary antibodies: 1:2000; 1:4000; 1:7000 were unable to increase the specificity. In addition western blot performed excluding primary antibodies demonstrated that several of the unspecific bands were due to unspecific binding of the secondary antibody (Figure 31).

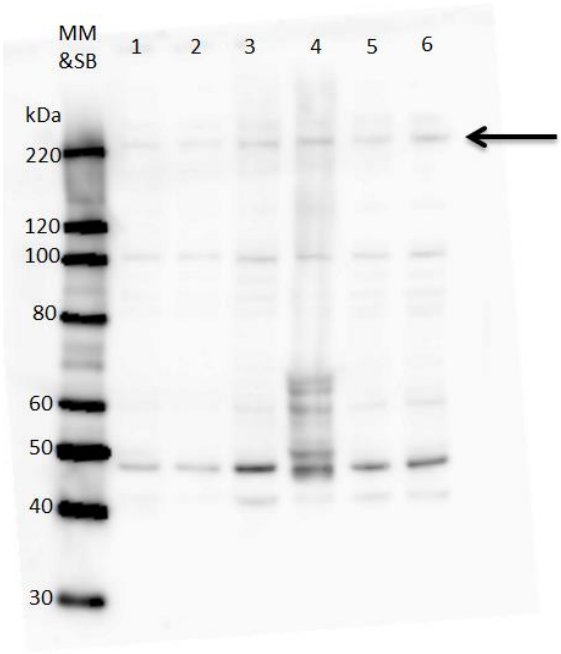



**Figure 30. Western blot analysis of cell lysates from A375 transfected cells using BRCA1 (D20) specific primary antibodies (1:1000). Secondary antibody was IgG Chicken-anti rabbit – HRP (1:40.000). MM= MagicMark™XP Western Protein Standard (Life Technologies). (1) negative control, (2) cells treated with only fugene, (3) cells transfected with plasmid without insert (pcDNA 3.1), (4) cells transfected with HA-tagged *BRCA1*, (5) cells transfected with plasmid without insert (pcDNA 3.1) and (6) cells transfected with HA-tagged *BRCA1*. In lane 5 and 6 cell lysates from cells harvested approximately 48 hours after transfection were applied. In the other lanes cell lysate which were harvested approximately 24 hours after transfection were used. Arrow is indicating expected band size. For each lane, 4.46 µg proteins were added.**

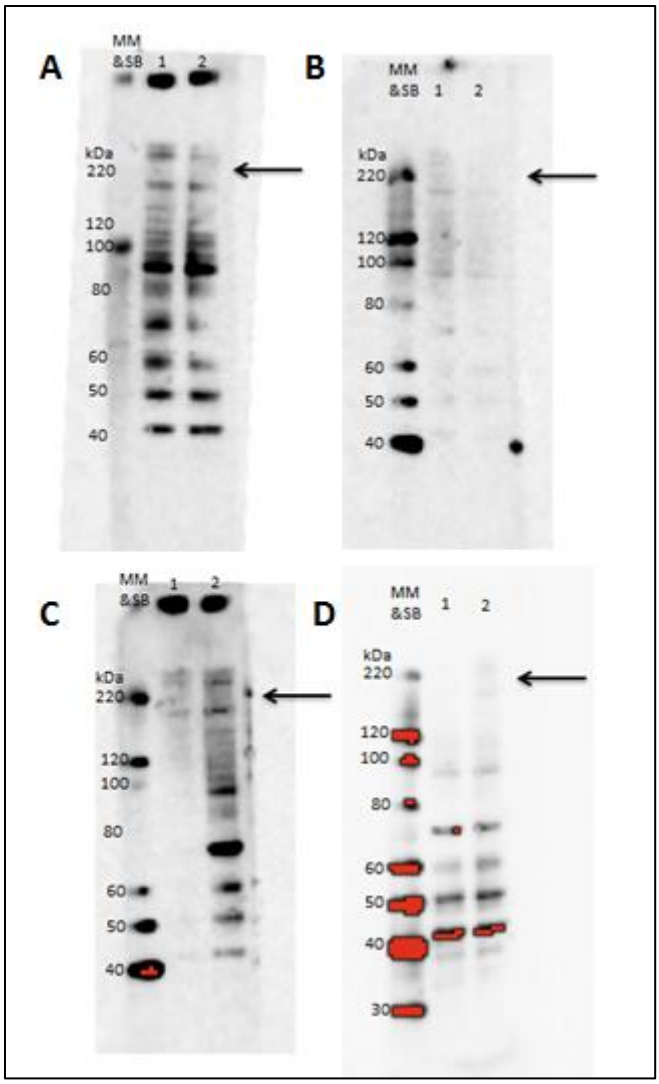



**Figure 31. Western blot analysis of cell lysates from HEK-293 cells. HA specific primary antibodies in different concentrations and the secondary antibody IgG Chicken-anti rabbit – HRP (1:40.000) were used. MM= MagicMark™XP Western Protein Standard (Life Technologies) and SB= SeeBlue Plus2 Pre-Stained Standard (Life Technologies). (1) Cells transfected with plasmid without insert (pcDNA 3.1) and (2) cells transfected with HA-tagged *BRCA1*. Cells were harvested approximately 24 hours after transfection. Arrow is indicating expected band size. (A) Only secondary antibodies were used, (B) primary antibody concentration 1:7000, (C) primary antibody concentration 1:4000 and (D) primary antibody concentration 1:2000. For each lane, 4.46 µg proteins were added.**

# 5. Discussion

## 5.1. PALB2

### 5.1.1 M13 method development

In order to screen multiple exons in multiple patients for possible disease causing mutations, it was necessary to develop a more efficient screening method. M13 based sequencing is just that, a more efficient sequencing method. M13 based sequencing allowed for the use of only one forward and one reverse primer in the sequencing reaction mixture. This decreases sequencing mixtures to only two, not including template DNA. This is not only more efficient, but limits the risk for possible pipetting errors when several different exons need to be studied. Several PCR conditions including different concentrations of template in sequencing mixture, two different polymerases used in PCR amplification and three (counting primers from kit) different M13 primers from different providers and purified differently were tested. Finally, only the quality of the M13 primers was shown to be of importance. The initial M13 primers purchased from Sigma-Aldrich were purified by desalting, while M13 primers purchased from Life Technologies were purified by desalting and gel filtration. In addition, the latter primers were assayed for function in automated sequencing prior to the distribution to the consumer. According to our results it seems to be important that M13 primers are both desalted and that unfinished oligonucleotides from a different length are removed before satisfactory sequencing can be attained.

### 5.1.2 Screening for PALB2 mutations

In total 43 patients were screened for HBOC causal mutations in *PALB2*. However, no obvious deleterious mutations, such as nonsense mutations or deletion/insertions, generating a frameshift during translation were identified in *PALB2*. Nevertheless, in these 43 patients 14 *PALB2* sequence variants were identified, 6 were found in exons, 7 in introns and one in the 5'UTR (Table 15). These 14 variants were subsequently evaluated and classified as either benign/likely benign, VUS or probably pathogenic/pathogenic based on *in silico* analysis.

Nine variants were classified as benign or likely benign (Table 15): 4 exonic and 5 intronic. One of them, c.3114-51A is the reported ancestral allele according to dbSNP (rs249936), and all 43 patients investigated were homozygous for this variant. The reference sequence NM_024675.3, which was used for comparison uses the minor allele at this position, with MAF of 0.000. Another four variants are placed in this group based on MAF, which for all of them is above 0.146 including c.1676A>G, c.2586+58C>T, c.2996+264T>C and 3201+101A>G (rs249954, rs420259 and rs249935, respectively). However, c.2586+58C>T have been reported to increase breast cancer risk with 36% in a Chinese population by Chen *et al.* (2008) [70]. The study reporting an increased cancer risk for

c.2586+58C>T have done a control study of common *PALB2* polymorphisms. They did a statistical analysis of 997 individuals with breast cancer and 1,008 individuals as controls to see if there is a correlation with increased breast cancer risk. This comparison revealed the slight increase in breast cancer risk for c.2586+58C>T [70]. Subsequently, c.2586+58C>T are classified as likely benign compared to the other variants with MAF above 0.146 classified as benign. Three exonic variants were classified as likely benign, c.1010C>T and c.2014G>C based on prediction programs which classified them as benign (except for SIFT), MAF above 0.01 indicating they are polymorphisms (0.012 and 0.014, respectively) and frequency in normal controls [63, 67]. Erkko and colleagues (2007) found c.1010T>C to be more common in the normal controls than individuals with familial breast cancer (10% compared to 8%, respectively) in a Finnish population [63]. Silvestri and colleagues (2010) found c.2014G>C in 10% of controls compared to 11.3% non-*BRCA1/2* male breast cancer cases in Central-Italy [67]. The third, c.3300T>G is a synonymous mutation with a MAF of 0.018. This variant did not seem to affect splicing, it is located outside known functional domains (see Figure 6) and are found to be common in controls [15, 67]. Silvestri and colleagues (2010) identified also this variant during screening of male breast cancer cases and this variant was also identified as normal in controls (12.2% in controls compared to 12.3% of breast cancer cases) [67]. Rahman and colleagues (2007) investigated non-*BRCA1/2* breast cancer patients in the United Kingdoms, and found c.3300T>G to be as frequent in controls as in non-*BRCA1/2* breast cancer patients (3% for both) [15]. The last mentioned identified variant in this group, c.212-58A>C is a polymorphism (MAF 0.014) that do not seem to impact splicing based on prediction programs and it is found to be common in controls [67, 69]. Both Silvestri and colleagues (2010) and Dansonka-Mieszkowska and colleagues (2010) have found this variant to be common in controls. The latter did a study based on variants found in ovarian carcinomas and further investigated them in a larger Polish population (sporadic, familial and controls) [69].

Five variants in *PALB2* were classified as VUS (Table 15): 2 intronic, 2 exonic and one located in the 5'UTR. Of the two identified intronic variants, c.2748+92T>G has not been earlier identified (to our knowledge). The SSF-like prediction program predicts a 5'ss (score 92.2) at the mutation site for c.2748+92G, but none of the other splice-prediction programs predict a 5'ss here. Interestingly this variant was identified in one patient diagnosed with breast cancer (patient 7) in which no other sequence variants were identified in *PALB2*. In the other patient (patient 9) with this sequence variant (c.2748+92T>G) also a missense mutation c.1676A>G and another intronic variant c.3201+101A>G was identified. The other intronic variant, c. 2749-18C>T have been identified by Adank and colleagues (2011) during their screening for *PALB2* mutations in non-*BRCA1/2* patients in Netherlands [64]. They found no indication for splice defects using Alamut version 1.5 and neither did we by using Alamut version 2.2. Further studies including these two variants are necessary to reveal the possible significance of these variants. The two exonic variants, c.2794G>A and c.2993G>A were

both predicted by PolyPhen2 and SIFT to be deleterious, while AlignGVGD predicts both to be benign. MutationTaster predicts c.2794G>A to be a polymorphism while c.2993G>A to be disease causing. The sequence variant c.2794G>A causes the replacement of a valine with a methionine at position 932 in the PALB2 protein. Although both amino acids are hydrophobic, valine is conserved in 6 of the 9 species investigated during multiple sequence alignment at this position (see Figure 19). In addition this amino acid is located in the WD40 (aa 853-1186) domain of PALB2 a domain important for binding and subsequent localization of the BRCA2 protein at DSBs and binding of RAD51. It remains to be seen if this amino substitution is important for the function of PALB2. Prediction programs indicates that this variant might produce a new 5'ss 2 nucleotides downstream of mutation. SpliceSiteFinder-like, MES and HSF predicts a 5'ss at this position. Subsequently, mRNA of the c.2794G>A should be analyzed to rule out possible missplicing for this variant. The other exonic variant classified as a VUS, c.2993G>A causes the replacement of the tiny amino acid glycine with glutamine at position 988 in the PALB2 protein. Glycine is highly conserved at this position and located in the WD domain. The change of the hydrophobic glycine with the polar glutamine might have an impact on protein folding and thereby function. However, although there are several indications that these two variants may influence the structure and function of PALB2, they have also been identified in normal controls. Rhaman and colleagues (2010) have found them to be as normal in controls as in individuals with HBOC when they investigated 1846 individuals with HBOC and 2168 controls: 2% compared to 1.8%, respectively for c.2993G>A and 0.3% for both regarding c.2794G>A [15]. Errko and colleagues (2007) found c.2794G>A to be more frequent in controls (5.1% of 315 in controls compared to 1.8% of 113 in patients) [63]. Also Kuusisto and collaborators reported an occurrence of this sequence variant in 3.8% (of 372) in a normal Finnish population compared to 1.2% (of 82) in the investigated HBOC population [66].

None of the observed variants were either nonsense mutations or frame shift mutations which are obvious deleterious mutations. Exclusively missense mutations were identified in our patient material. The patients investigated in this study for *PALB2* mutations were previously shown to be negative for mutations in *BRCA1* and *BRCA2*. They were therefor strong candidates for mutations in *PALB2* according to Poumpouridou (2011). He claimed that *PALB2* mutations have been found in most populations investigated so far [12]. This includes Finnish, Australian, Spanish and individuals from USA, amongst others. However, common for all these studies are the low frequencies of pathogenic *PALB2* mutations, ranging from 1-4% [63, 98-101]. In this current study, 43 patients were screened and no obvious pathogenic mutations were found in this population. Accordingly, we can conclude that mutations in *PALB2* are not a very common cause of HBOC in the northern Norwegian HBOC population. However, larger cancer population studies are needed. It would have been interesting to screen a lot of patients at the same time and perhaps some genes coding for other proteins involved in the HRR-pathway, using next generation sequencing.

## 5.2. BRCA1 and BRCA2

To evaluate VUS from individuals of families with breast and ovarian cancer, 20 different *BRCA1* VUS and 15 different *BRCA2* VUS were investigated. They were evaluated for their influence on protein structure and function with *in silico* tools. In addition these variants were evaluated for their effect on splicing and to investigate if both alleles are expressed in blood. The latter being an indication of whether there are other mutations not previously identified, that might explain the absence of a particular allele.

We isolated RNA from whole blood collected from individuals in HBOC families. RNA was subsequently reverse transcribes to cDNA and quality tested by the use of an APRT-test (described in methods). Twenty *BRCA1* variants and 15 *BRCA2* variants were investigated, corresponding to 19 individuals with *BRCA1* variants and 18 individuals with *BRCA2* variants. All samples were checked for the presence of both alleles and for a different band pattern of the investigated cDNA compared with cDNA from the normal control. Samples with different band pattern after PCR and gel electrophoresis were subsequently sequenced and investigated for aberrant splicing pattern.

When we investigated if both alleles were transcribed, intronic variants were more challenging to verify due to out-splicing of introns (except for *BRCA1* c.213-5T>A, discussed later). The variants we were not able to verify includes c.-33-29delAAAAA and c.548-17G>T from *BRCA1* and c.68-7T>A from *BRCA2*. For these variants, polymorphisms located elsewhere could be used as a marker for further investigation of non-identified mutations affecting transcription. Although both alleles were identified for the other samples, amount was not measured. This would be possible with the use of quantitativePCR (qPCR).

Investigation of aberrant splicing in the 36 individuals using gel electrophoresis for size separation of amplified cDNA revealed two potential *BRCA1* variants causing aberrant splicing (se gel images Figure 21 and Figure 23). These variants, *BRCA1* c.213-5T>A and c.5434C>G were subsequently sequenced, together with normal control for possible identification of aberrant splicing. The *BRCA1* c.213-5T>A variant produce transcripts with inclusion of 59 nt of the 3' end of intron 5. This variant is not near the new 3'ss, but might weaken the original enough to activate the cryptic splice site 59 nt upstream. Inclusion of 59 nt of intron 5 results in a fram-shift that results in a stop codon after 75 codons. This inclusion have been reported before, but then caused by another mutation, c.213-11T>G [102]. *BRCA1* c.5434C>G leads to exclusion of exon 23, corresponding to 61 nt and thereby result in a frame-shift. This variant has been previously reported by Gaildrat *et al.*(2010). In their study they used a mini-gene analysis and found this variant to exclude exon 23 from most transcripts [87].

Gel electrophoresis for size separation of amplified cDNA revealed at first one *BRCA2* variant potentially causing aberrant splicing (se gel image Figure 27), c.9116C>T. Llort *et al.* (2002) have classified this as deleterious simply based on location, in exon 23 only two nucleotides upstream from the 5'ss [93]. Evaluation of splicing predictions *in silico* indicates a possible pathogenic effect by weakening the 5'ss. We had two different individuals carrying this variant, cDNA synthesis and analysis by gel electrophoresis exhibit different band patterns for both individuals compared to the control. However, we were not able to achieve a clear indication of aberrant splicing from the subsequent sequencing data obtained of these aberrant PCR products. It would be interesting to isolate the different PCR products and sequence these individually in order to get a better understanding of the mechanism induced by this sequence. Another *BRCA2* variant, c. 68-7T>A did not have a different band pattern from the normal control, but upon evaluation of sequences skipping of exon 3 was identified. This variant is located in intron 2, close to the 3'ss. Sanz *et al* (2010) have studied this variant previously using a mini-gene analysis. They showed that this variant cause partial exon 3 skipping [89]. However, Théry *et al.* (2011) was not able to reproduce the results from the mini-gene assay [88]. Yet another study, Vreeswijk *et al.* (2008), identified increased skipping for this variant. They tested mRNA from fibroblast of patients in high-risk breast cancer families [103]. Exon 3 is 149 nt long, resulting in an in-frame skipping. Since this skipping is in-frame and also found in the normal control, it still classified as a VUS. Even though, exon 3 is responsible for coding the PALB2 binding domain (aa 21-39) in BRCA2 (Figure 4). Since PALB2 is responsible for locating BRCA2 at the DSB, one might argue that a high proportion of exon 3 skipping transcripts might cause an insufficient amount of BRCA2 protein with the PALB2 binding domain and may accordingly lead to an increased risk of developing breast and ovarian cancer. Whether or not this variant produce a high enough proportion is difficult to say out from our experiments. Further research, including analysis of co-segregation of this sequence variant in the family, the normal population frequency studies, and loss of heterozygousity studies in tumors will be needed to fully understand the consequence of this variant.

Transcription of *BRCA1* and *BRCA2* generated several alternative spliced transcripts [94-97]. Accordingly, during the current study we were able to identify four alternative spliced transcripts in blood samples of HBOC families as well as in the normal control. A literature search classified all four transcripts as normal alternative splicing variants. Of the four normal transcript variants, three were found in *BRCA1* and one in *BRCA2* (including exon 3 skipping as discussed earlier). In *BRCA1*, we found an alternative transcript which skips the last 6 nucleotides of exon 1 (ΔGTAAAG), a second that skips the first 3 nucleotides in exon 8 (ΔCAG) and a third one skipping both exon 9 and exon 10 (Δ9,10). All 4 variants were summarized in Table 18. The ΔGTAAAG in exon 1 was first reported by Fetzer *et al.* (1998), where they identified this transcript in lymphocytes [94]. Later, both ΔGTAAAG in exon 1 and ΔCAG in exon 8 were reported by Jakubowska *et al.* (2001) who investigated transcripts present in lymphocytes [95]. Jakubowska and collaborators were the first to report a transcript

skipping the first 3 nucleotides (ΔCAG) in exon 8, as well as in exon 14. Skipping of ΔCAG in exon 14 was not possible to detect in this current study based on primer positions. Another alternative transcript Δ9,10 is classified as a normal transcript variant since this transcript has been identified in non-malignant breast-cells [96]. Skipping of both exon 9 and 10 results in a deletion of 41 amino acids in a region of BRCA1 outside of the currently identified functional domains.

The presence of normal alternative transcripts complicates analysis of splice variants, since it might be difficult to see what is natural alternative splicing and what is caused by the variant investigated. In addition, the mutations investigated might even produce an increase in alternatively spliced transcripts and may accordingly be pathogenic. One way to investigate this would be by using qPCR, where the amounts of alternative transcripts can be compared within a sample and between HBOC samples and normal controls. Another thing to keep in mind in studying the alternative transcripts is that they might differ in amount between different tissue. In this current study we use exclusively cDNA generated from RNA isolated from peripheral lymphocytes. How the amount of the different transcript variants vary in other tissues is currently not systematically studied to our knowledge.

After studying 19 individuals with variants in *BRCA1* and 18 individuals with variants in *BRCA2* by *in silico* evaluation, splicing analysis and test for both alleles as a marker for other undetected variants, a lot of variants were still classified as VUS. The variants classified as VUS will first be discussed, followed by variants classified as benign/likely benign and thereafter variants classified as pathogenic.

The VUS includes 6 *BRCA1* variants and 6 *BRCA2 variants* (listed in Table 16 and Table 17). Of these 12 currently studied variants, 4 *BRCA1* variants, c.-33-29delAAAAA, c.5075A>C, c.5125G>A and c.5513T>G and 1 *BRCA2* variant, c.8323A>G were novel. The lack of information and divergent prediction results (see Table 23 and

Table 24) together with the lack of impact on splicing lead us to conclude that these variants should be classified as VUS. Two other variant classified as a VUS are *BRCA2* c.68-7T>A, which increase skipping of exon 3 and c.9116C>T which we were not able to confirm expected aberrant splicing, as previously discussed. Another two *BRCA2* variants, c.750G>A and c.4068G>A were synonymous mutations. Analysis of the sequence of the latter was not possible due to weak PCR-products. *BRCA2* c.750G>A have not been previously reported in the literature to our knowledge, neither have *BRCA2* c.7301A>C. The general lack of information of frequency, reports in literature and functional studies lead us to conclude that also these variants should be classified as VUS.

Variants classified as benign or likely benign includes 7 *BRCA1* variants, c.548-17G>T, c1487G>A, c.2521C>T, c.3418A>G, c.5117G>C, c.486G>T and c.1419C>T and 7 *BRCA2* variants, c.40A>G, c.4828G>A, c.5272_5274delAAT, c.6821G>T, c.2680G>A, c.3568C>T and c.6100C>T. One of the *BRCA2* variants, c.40A>G also showed skipping of exon 3, but since this is as mentioned found in the normal control and had no indication of increased skipping. Prediction programs in Alamut predicts the formation of a 3'ss, but this splicing was not detected in cDNA sample. Based on this it is probably not pathogenic, although it could be included if samples were tested with qPCR. Another *BRCA2* variant classified as likely benign is c.5272_5274delAAT. This variant is novel, but does not seem to produce alternative splicing, or affect protein since it is an in-frame deletion and between to BRC repeats. Two *BRCA1* variants, c.486G>T and c.1419C>T are both synonymous and gave no indication of alternative transcripts and both alleles were transcribed. The remaining benign/likely benign variants include four *BRCA1* variants c.548-17G>T, c.1487G>A, c.2521C>T and c.3418A>G as benign and four *BRCA2* variants, c.4828G>A, c.6821G>T, c.3568C>T and c.6100C>T as benign/likely benign. Classification of these variants are based on literature, predictions and splicing/transcript analysis (se Table 16, Table 17, Table 23 and

Table 24). Lindor *et al.* has evaluated several of these variants, by using a quantitative "posterior probability model" for assessing VUS in *BRCA1* and *BRCA2*. The model is based on multiple factors like personal and family history, co-segregation with disease, co-occurrence of pathogenic variants in "trans" and pathology profile [72]. Variants that they have classified as benign/likely benign included in our current study includes *BRCA1* c. 548-17G>T, c.1487G>A, c.2521C>T, c.3418A>G and c.5117G>C and *BRCA2* c.2680G>A, c.3568C>T, c.6100C>T and c.6821G>T.

In the last group of the variants classified during this study were variants which are either pathogenic or probably pathogenic. We were able to classify six variants as either pathogenic or probably pathogenic based on *in silico* investigation of variants and splicing analysis (Table 16 and Table 17). The *BRCA1* variant c.1A>G has been shown earlier to reduce the N-terminal end of the protein by 17 aa, due to the use of an alternative start codon further downstream. Liu *et al.* (2000) proved this in a cell-system by transfecting cells with WT-*BRCA1* and mutant-*BRCA1* in plasmids. However, the cells transfected with WT *BRCA1* produced transcripts generated from both the first translation start codon (AUG), but also transcripts which used the translational start 17 codons further downstream. Cells transfected with mutant only used the second translation start, 17 codons downstream from the first AUG [33]. Liu *et al.* (2000) also investigated if this alternative translation initiation codon from thymus homogenate, and found when analyzing the N-terminal protein sequence that approximately 20% of the translations started at the second AUG [33]. The lack of the first 17 aa might impact the RING finger domain, and thereby inhibit BARD1 association and interaction with BAP1 and E2F transcription factors [27]. Two other mutations also identified in our patient material, *BRCA1* c.130T>A (Cys44Ser) and *BRCA1* c.140G>T (Cys47Phe) are also located in the RING-finger domain.

Experimental altering of the Cysteine in position 44 was shown to result in the formation of a BRCA1 RING-domain that failed to fold properly [74]. In addition, Millot *et al*. (2011) investigated these two variants by the use of a "yeast localization phenotype assay", which is based on the accumulation of BRCA1 in a single inclusion body in the yeast nucleus. They discovered that the two variants partially abrogate the BRCA1 inclusion body formation seen for WT BRCA1 [75]. Two variants classified as pathogenic have already been mentioned in regards to splicing, *BRCA1* c.213-5T>A and *BRCA2* c.5434C>G. Another *BRCA1* variant, c.5123C>T (Ala1708Val) is classified as probably pathogenic based on the substitution of the small alanine with the much larger valine, its location in the BRCT domain, prediction programs and protein modelling that indicates impact on folding [86]. In addition *BRCA1* c.5123C>T have been shown to have intermediate transactivation activity in the expressed BRCA1 protein, while still being able to form foci when stimulated [85]. Based on this information, Lovelock *et al.* (2007) suggests that this variant might be associated with low or moderate cancer risk [85]. Of the variants located in *BRCA2* c.7878G>C is located in a sub-helical domain in the DSS1 binding domain [104] and the amino acid at this position is highly conserved. Based on co-segregation, personal and family history, co-occurrence in "trans" and pathology profile mentioned before taken in consideration by Lindor *et al.* (2011), this variant was classified as pathogenic [72]. The other *BRCA2* variant, c.8177A>G produce a BRCA1 protein with an aa-substitution located in the OB-domain (see Figure 4) which is known to associate with DSS1 [28] –and references within. Both tryptophan at position 2726 and adenine at position c.8177 are highly conserved, and all prediction programs in Alamut classifies this as pathogenic which results in the classification of this as probably pathogenic.

It would be interesting to investigate *BRCA1* c.130T>A and *BRCA1* c.140G>T's ability to form foci in the presence of DSB in BRCA1-negative cells. Or the variants located in the BRCT trans activating domain's ability to still activate transcription, from *BRCA1* c.5075A>C to c.5513T>G listed in Table 16 in results.

## 5.3. Cell work and western blot for further investigation of the VUS

To investigate the consequences of some of the mutations found in our patient material at protein level, we did some preliminary studies to express mutant BRCA1 protein in *in vitro* cell systems and visualization of BRCA1 with western blot analysis. For this assay, the available plasmid pcDNA3 HA-*BRCA1* 1387Ala was successfully reversed to pcDNA3 HA-*BRCA1* WT. At the same time, the pcDNA3 HA-*BRCA1* 1387Ala was used in transfection studies both of A375 and HEK-293 cells. These are cells known to be easily transfected, and therefore used during optimization for western blot

analysis. Western blot was intended to be used to study BRCA1 expression in transfected cells. During our initial studies however we were unable to detect BRCA1 protein in cell lysates from A375 transfected cells (Figure 30). In this first test we used BRCA1 (D-20) specific antibodies. Subsequently we used HA-specific antibodies, since the transfected pcDNA3 HA-*BRCA1* 1387Ala is HA-tagged. However, neither these antibodies were able to show us BRCA1 specific protein bands compared to control. Different concentrations of primary antibodies were tested (1:1000-7000) to see if we could increase specificity, unfortunately without success (Figure 31). We were also able to show that our secondary antibody demonstrated several unspecific protein bands (Figure 31). Currently, we were not able to distinguish if transfection rates were low since the transfected plasmid was 12kb large or if levels of BRCA1 expression were below the detection range with the used BRCA1 antibodies. Accordingly, further studies including studying different primary antibodies are needed to optimize western blot analysis before the influence of variants of unknown significance can be studies using this technology.

# 6. Concluding remarks

M13 based sequencing is an efficient way to sequence several different exons and samples at the same time. Using this newly introduced method, no real causal *PALB2* pathogenic mutations were discovered screening a limited HBOC population. Applying new technology, such as next generation sequencing (NGS) will allow us in the future to investigate appropriate larger amount of patients in lesser time. In addition this NGS technology will allow us to investigate additional other genes coding for proteins involved in HRR.

In our study on the identified variants of unknown clinical significance identified in *BRCA1* and *BRCA2*, only 3 variants (*BRCA1* c.213-5T>A, *BRCA1* c.5434C>G and *BRCA2* c.68-7T>A) were shown to cause aberrant splicing. *BRCA1* c.213-5T>A has not been previously described. However, another mutation 6 nt upstream, *BRCA1* c.213-11T>G causes the same aberrant splicing resulting in inclusion of 59 nt of intron 5 in the modified transcript. *BRCA1* c.5434C>G has been reported before, and we were able to confirm the earlier published results. The *BRCA2* c.68-7T>A is still classified as a VUS. All of the exonic variants (except *BRCA2* c.3568C>T and *BRCA2* c.4068G>A due to amplification problems) in *BRCA1/BRCA2* included in this study were transcribed from both alleles when assessed studying RNA produced in whole blood. Accordingly, no additional variants were disclosed

As mentioned, further investigation of some of these variants on protein level would be interesting. An assay is under development, and will hopefully be completed during the summer months. The plan is

to test different mutations using BRCA1-negative cells, and see if transfected mutants can complement the cells lack of functional BRCA1 in these cells.

## References


1. *Kreftregisteret*. 2013 [cited 2013; Available from: www.Kreftregisteret.no.
2. Knudson, A.G., Jr., *Mutation and cancer: statistical study of retinoblastoma.* Proc Natl Acad Sci U S A, 1971. **68**(4): p. 820-3.
3. Berger, A.H., A.G. Knudson, and P.P. Pandolfi, *A continuum model for tumour suppression.* Nature, 2011. **476**(7359): p. 163-9.
4. *unn.no*. [cited 2013; Available from: http://www.unn.no/kompetansesenter-for-arvelig-kreft/category26897.html.
5. Miki, Y., et al., *A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.* Science, 1994. **266**(5182): p. 66-71.
6. Wooster, R., et al., *Identification of the breast cancer susceptibility gene BRCA2.* Nature, 1995. **378**(6559): p. 789-92.
7. Olopade, O.I., et al., *Advances in breast cancer: pathways to personalized medicine.* Clin Cancer Res, 2008. **14**(24): p. 7988-99.
8. *HGMDp*. 2013 [cited 2012, 15. May; Available from: https://portal.biobase-international.com/hgmd/pro/start.php.
9. *BIC database*. 2013 [cited 2013, 15.May; Available from: http://research.nhgri.nih.gov/projects/bic/Member/index.shtml.
10. Birch, J.M., et al., *Prevalence and diversity of constitutional mutations in the p53 gene among 21 Li-Fraumeni families.* Cancer Res, 1994. **54**(5): p. 1298-304.
11. Martin, A.M. and B.L. Weber, *Genetic and hormonal risk factors in breast cancer.* J Natl Cancer Inst, 2000. **92**(14): p. 1126-35.
12. Poumpouridou, N. and C. Kroupis, *Hereditary breast cancer: beyond BRCA genetic analysis; PALB2 emerges.* Clin Chem Lab Med, 2012. **50**(3): p. 423-34.
13. Kee, Y. and A.D. D'Andrea, *Expanded roles of the Fanconi anemia pathway in preserving genomic stability.* Genes Dev, 2010. **24**(16): p. 1680-94.
14. Seal, S., et al., *Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles.* Nat Genet, 2006. **38**(11): p. 1239-41.
15. Rahman, N., et al., *PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene.* Nat Genet, 2007. **39**(2): p. 165-7.
16. Meindl, A., et al., *Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene.* Nat Genet, 2010. **42**(5): p. 410-4.
17. Johnson, J., et al., *Mutation analysis of RAD51L1 (RAD51B/REC2) in multiple-case, non-BRCA1/2 breast cancer families.* Breast Cancer Res Treat, 2011. **129**(1): p. 255-63.
18. Loveday, C., et al., *Germline mutations in RAD51D confer susceptibility to ovarian cancer.* Nat Genet, 2011. **43**(9): p. 879-82.
19. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
20. Alberts, B., *Molecular biology of the cell*. 5th ed. 2008, New York: Garland Science.
21. Bekker-Jensen, S. and N. Mailand, *Assembly and function of DNA double-strand break repair foci in mammalian cells.* DNA Repair (Amst), 2010. **9**(12): p. 1219-28.
22. Caestecker, K.W. and G.R. Van de Walle, *The role of BRCA1 in DNA double-strand repair: past and present.* Exp Cell Res, 2013. **319**(5): p. 575-87.
23. Weinberg, R.A., *The biology of cancer*. 2007, New York ; Abingdon: Garland Science. 864 p. in various pagings.
24. Hall, J.M., et al., *Linkage of early-onset familial breast cancer to chromosome 17q21.* Science, 1990. **250**(4988): p. 1684-9.

25. Chen, Y., et al., *BRCA1 is a 220-kDa nuclear phosphoprotein that is expressed and phosphorylated in a cell cycle-dependent manner.* Cancer Res, 1996. **56**(14): p. 3168-72.
26. Meza, J.E., et al., *Mapping the functional domains of BRCA1. Interaction of the ring finger domains of BRCA1 and BARD1.* J Biol Chem, 1999. **274**(9): p. 5659-65.
27. Narod, S.A. and W.D. Foulkes, *BRCA1 and BRCA2: 1994 and beyond.* Nat Rev Cancer, 2004. **4**(9): p. 665-76.
28. Roy, R., J. Chun, and S.N. Powell, *BRCA1 and BRCA2: different roles in a common pathway of genome protection.* Nat Rev Cancer, 2012. **12**(1): p. 68-78.
29. Zhang, J., et al., *Chk2 phosphorylation of BRCA1 regulates DNA double-strand break repair.* Mol Cell Biol, 2004. **24**(2): p. 708-18.
30. Sy, S.M., M.S. Huen, and J. Chen, *PALB2 is an integral component of the BRCA complex required for homologous recombination repair.* Proc Natl Acad Sci U S A, 2009. **106**(17): p. 7155-60.
31. Xu, B., et al., *Phosphorylation of serine 1387 in Brca1 is specifically required for the Atm-mediated S-phase checkpoint after ionizing irradiation.* Cancer Res, 2002. **62**(16): p. 4588-91.
32. Cortez, D., et al., *Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks.* Science, 1999. **286**(5442): p. 1162-6.
33. Liu, J., et al., *Initiation of translation from a downstream in-frame AUG codon on BRCA1 can generate the novel isoform protein DeltaBRCA1(17aa).* Oncogene, 2000. **19**(23): p. 2767-73.
34. *BRCA1 in www.ensembl.org*. 2013 [cited 2013 15. April]; Available from: http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000012048;r=17:41196312-41322290.
35. Wooster, R., et al., *Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13.* Science, 1994. **265**(5181): p. 2088-90.
36. Hondow, H.L., et al., *A high-throughput protocol for mutation scanning of the BRCA1 and BRCA2 genes.* BMC Cancer, 2011. **11**: p. 265.
37. Davies, A.A., et al., *Role of BRCA2 in control of the RAD51 recombination and DNA repair protein.* Mol Cell, 2001. **7**(2): p. 273-82.
38. Li, J., et al., *DSS1 is required for the stability of BRCA2.* Oncogene, 2006. **25**(8): p. 1186-94.
39. *SHFM1 at www.ncbi.nlm.nih.gov*. [cited 2013 15. April]; Available from: www.ncbi.nlm.nih.gov/gene/7979.
40. *BRCA2 in www.ensemble.org*. [cited 2013 15. April]; Available from: http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000139618;r=13:32889611-32973805.
41. Xia, B., et al., *Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2.* Mol Cell, 2006. **22**(6): p. 719-29.
42. Bleuyard, J.Y., et al., *ChAM, a novel motif that mediates PALB2 intrinsic chromatin binding and facilitates DNA repair.* EMBO Rep, 2012. **13**(2): p. 135-41.
43. Buisson, R. and J.Y. Masson, *PALB2 self-interaction controls homologous recombination.* Nucleic Acids Res, 2012. **40**(20): p. 10312-23.
44. Tazi, J., N. Bakkour, and S. Stamm, *Alternative splicing and disease.* Biochim Biophys Acta, 2009. **1792**(1): p. 14-26.
45. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
46. Lewin, B., et al., *Lewin's genes X*. 10th ed. 2011, Sudbury, Mass.: Jones and Bartlett. xxvi, 930 p.

47. Padgett, R.A., *New connections between splicing and human disease.* Trends Genet, 2012. **28**(4): p. 147-54.
48. Andresen, B.S., Krainer, A. R., *When the genetic code is not enough – How sequence variations can affect pre-mRNA splicing and cause (complex) disease*, in *Chapter 15 (pp.165-182) in Genetics of Complex Human Diseases*, L. Almasy, A. Al-Chalabi, Editor. 2009, New York, USA: Cold Spring Harbor Laboratory Press.
49. Ast, G., *How did alternative splicing evolve?* Nat Rev Genet, 2004. **5**(10): p. 773-82.
50. Kelemen, O., et al., *Function of alternative splicing.* Gene, 2013. **514**(1): p. 1-30.
51. Graveley, B.R., *Sorting out the complexity of SR protein functions.* RNA, 2000. **6**(9): p. 1197-211.
52. Cooper, T.A. and W. Mattox, *The regulation of splice-site selection, and its role in human disease.* Am J Hum Genet, 1997. **61**(2): p. 259-66.
53. Roca, X., et al., *Features of 5'-splice-site efficiency derived from disease-causing mutations and comparative genomics.* Genome Res, 2008. **18**(1): p. 77-87.
54. Qiagen. *EZ1 DNA Handbook*. 2004 [cited 09.04.2013; Available from: http://projects.nfstc.org/workshops/resources/articles/EZ1%20DNA%20Handbook.pdf .
55. *PAXgene Blood RNA Kit Handbook v.2*, QIAGEN.
56. Ng, P.C. and S. Henikoff, *Predicting deleterious amino acid substitutions.* Genome Res, 2001. **11**(5): p. 863-74.
57. Kumar, P., S. Henikoff, and P.C. Ng, *Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm.* Nat Protoc, 2009. **4**(7): p. 1073-81.
58. Tavtigian, S.V., et al., *Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral.* J Med Genet, 2006. **43**(4): p. 295-305.
59. Mathe, E., et al., *Computational approaches for predicting the biological effect of p53 missense mutations: a comparison of three sequence analysis based methods.* Nucleic Acids Res, 2006. **34**(5): p. 1317-25.
60. Schwarz, J.M., et al., *MutationTaster evaluates disease-causing potential of sequence alterations.* Nat Methods, 2010. **7**(8): p. 575-6.
61. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
62. Houdayer, C., *In silico prediction of splice-affecting nucleotide variants.* Methods Mol Biol, 2011. **760**: p. 269-81.
63. Erkko, H., et al., *A recurrent mutation in PALB2 in Finnish cancer families.* Nature, 2007. **446**(7133): p. 316-9.
64. Adank, M.A., et al., *PALB2 analysis in BRCA2-like families.* Breast Cancer Res Treat, 2011. **127**(2): p. 357-62.
65. Blanco, A., et al., *Detection of a large rearrangement in PALB2 in Spanish breast cancer families with male breast cancer.* Breast Cancer Res Treat, 2012. **132**(1): p. 307-15.
66. Kuusisto, K.M., et al., *Screening for BRCA1, BRCA2, CHEK2, PALB2, BRIP1, RAD50, and CDH1 mutations in high-risk Finnish BRCA1/2-founder mutation-negative breast and/or ovarian cancer individuals.* Breast Cancer Res, 2011. **13**(1): p. R20.
67. Silvestri, V., et al., *PALB2 mutations in male breast cancer: a population-based study in Central Italy.* Breast Cancer Res Treat, 2010. **122**(1): p. 299-301.
68. Papi, L., et al., *A PALB2 germline mutation associated with hereditary breast cancer in Italy.* Fam Cancer, 2010. **9**(2): p. 181-5.

69. Dansonka-Mieszkowska, A., et al., *A novel germline PALB2 deletion in Polish breast and ovarian cancer patients.* BMC Med Genet, 2010. **11**: p. 20.
70. Chen, P., et al., *Association of common PALB2 polymorphisms with breast cancer risk: a case-control study.* Clin Cancer Res, 2008. **14**(18): p. 5931-7.
71. Frost, P., et al., *Complete mutation screening and haplotype characterization of the BRCA1 gene in 61 familial breast cancer patients from Norway.* Dis Markers, 2005. **21**(1): p. 29-36.
72. Lindor, N.M., et al., *A review of a multifactorial probability-based model for classification of BRCA1 and BRCA2 variants of uncertain significance (VUS).* Hum Mutat, 2012. **33**(1): p. 8-21.
73. Capanu, M., et al., *Assessment of rare BRCA1 and BRCA2 variants of unknown significance using hierarchical modeling.* Genet Epidemiol, 2011. **35**(5): p. 389-97.
74. Brzovic, P.S., et al., *Structure of a BRCA1-BARD1 heterodimeric RING-RING complex.* Nat Struct Biol, 2001. **8**(10): p. 833-7.
75. Millot, G.A., et al., *Assessment of human Nter and Cter BRCA1 mutations using growth and localization assays in yeast.* Hum Mutat, 2011. **32**(12): p. 1470-80.
76. Abkevich, V., et al., *Analysis of missense variation in human BRCA1 in the context of interspecific sequence variation.* J Med Genet, 2004. **41**(7): p. 492-507.
77. Begg, C.B., et al., *Variation of breast cancer risk among BRCA1/2 carriers.* JAMA, 2008. **299**(2): p. 194-201.
78. Sweet, K., et al., *Characterization of BRCA1 ring finger variants of uncertain significance.* Breast Cancer Res Treat, 2010. **119**(3): p. 737-43.
79. Bonnet, C., et al., *Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene.* J Med Genet, 2008. **45**(7): p. 438-46.
80. Goldgar, D.E., et al., *Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2.* Am J Hum Genet, 2004. **75**(4): p. 535-44.
81. Osorio, A., et al., *Classification of missense variants of unknown significance in BRCA1 based on clinical and tumor information.* Hum Mutat, 2007. **28**(5): p. 477-85.
82. Vallon-Christersson, J., et al., *Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families.* Hum Mol Genet, 2001. **10**(4): p. 353-60.
83. Laraqui, A., et al., *Mutation screening of the BRCA1 gene in early onset and familial breast/ovarian cancer in Moroccan population.* Int J Med Sci, 2013. **10**(1): p. 60-7.
84. Phelan, C.M., et al., *Classification of BRCA1 missense variants of unknown clinical significance.* J Med Genet, 2005. **42**(2): p. 138-46.
85. Lovelock, P.K., et al., *Identification of BRCA1 missense substitutions that confer partial functional activity: potential moderate risk variants?* Breast Cancer Res, 2007. **9**(6): p. R82.
86. Chenevix-Trench, G., et al., *Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance.* Cancer Res, 2006. **66**(4): p. 2019-27.
87. Gaildrat, P., et al., *The BRCA1 c.5434C->G (p.Pro1812Ala) variant induces a deleterious exon 23 skipping by affecting exonic splicing regulatory elements.* J Med Genet, 2010. **47**(6): p. 398-403.
88. Thery, J.C., et al., *Contribution of bioinformatics predictions and functional splicing assays to the interpretation of unclassified variants of the BRCA genes.* Eur J Hum Genet, 2011. **19**(10): p. 1052-8.

89. Sanz, D.J., et al., *A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients.* Clin Cancer Res, 2010. **16**(6): p. 1957-67.
90. Easton, D.F., et al., *A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes.* Am J Hum Genet, 2007. **81**(5): p. 873-83.
91. Stegel, V., et al., *The occurrence of germline BRCA1 and BRCA2 sequence alterations in Slovenian population.* BMC Med Genet, 2011. **12**: p. 9.
92. Simard, J., et al., *Evaluation of BRCA1 and BRCA2 mutation prevalence, risk prediction models and a multistep testing approach in French-Canadian families with high risk of breast and ovarian cancer.* J Med Genet, 2007. **44**(2): p. 107-21.
93. Llort, G., et al., *Low frequency of recurrent BRCA1 and BRCA2 mutations in Spain.* Hum Mutat, 2002. **19**(3): p. 307.
94. Fetzer, S., et al., *An alternative splice site junction in exon 1a of the BRCA1 gene.* Cancer Genet Cytogenet, 1998. **105**(1): p. 90-2.
95. Jakubowska, A., et al., *Detection of germline mutations in the BRCA1 gene by RNA-based sequencing.* Hum Mutat, 2001. **18**(2): p. 149-56.
96. Lu, M., et al., *Characterization of functional messenger RNA splice variants of BRCA1 expressed in nonmalignant and tumor-derived breast cells.* Cancer Res, 1996. **56**(20): p. 4578-81.
97. Muller, D., et al., *An entire exon 3 germ-line rearrangement in the BRCA2 gene: pathogenic relevance of exon 3 deletion in breast cancer predisposition.* BMC Med Genet, 2011. **12**: p. 121.
98. Wong, M.W., et al., *BRIP1, PALB2, and RAD51C mutation analysis reveals their relative importance as genetic susceptibility factors for breast cancer.* Breast Cancer Res Treat, 2011. **127**(3): p. 853-9.
99. Garcia, M.J., et al., *Analysis of FANCB and FANCN/PALB2 fanconi anemia genes in BRCA1/2-negative Spanish breast cancer families.* Breast Cancer Res Treat, 2009. **113**(3): p. 545-51.
100. Cao, A.Y., et al., *The prevalence of PALB2 germline mutations in BRCA1/BRCA2 negative Chinese women with early onset breast cancer or affected relatives.* Breast Cancer Res Treat, 2009. **114**(3): p. 457-62.
101. Casadei, S., et al., *Contribution of inherited mutations in the BRCA2-interacting protein PALB2 to familial breast cancer.* Cancer Res, 2011. **71**(6): p. 2222-9.
102. Friedman, L.S., et al., *Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families.* Nat Genet, 1994. **8**(4): p. 399-404.
103. Vreeswijk, M.P., et al., *Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs.* Hum Mutat, 2009. **30**(1): p. 107-14.
104. Yang, H., et al., *BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure.* Science, 2002. **297**(5588): p. 1837-48.
105. Xia, B., et al., *Fanconi anemia is associated with a defect in the BRCA2 partner PALB2.* Nat Genet, 2007. **39**(2): p. 159-61.

# Appendix A – samples and PCR programs

**Table 19. Patient samples used in *PALB2* screening. BC= Breast cancer, BOC= Breast and ovarian cancer, and OC = Ovarian cancer. "Diagnostic code" is cancer types found in the patients family; "cancer form" is the cancer form the patient has had; age = age at diagnosis.**

| Patient number in this study | Family number | Diagnostic code | Cancer form | Age |
|---|---|---|---|---|
| 1 | 1 | BC | BC | 63 |
| 2 | 2 | BOC | OC | 60 |
| 3 | 3 | BOC | OC | 56 |
| 4 | 4 | OC | OC | 50 |
| 5 | 5 | OC | OC | 59 |
| 6 | 6 | OC | OC | 37 |
| 7 | 7 | BC | BC | 57 |
| 8 | 8 | OC | OC | 52 |
| 9 | 9 | BC | BC | 53 |
| 10 | 10 | BC | BC | 69 |
| 11 | 11 | BC | BC | 38 |
| 12 | 12 | BC | BC | 49 |
| 13 | 7 | BC | BC | 61 |
| 14 | 13 | BC | BC | 57 |
| 15 | 13 | BC | BC | 58 |
| 16 | 14 | BC | BC | 48 |
| 17 | 14 | BC | BC | 57 |
| 18 | 15 | BC | BC | 51 |
| 19 | 16 | BC | BC | 59 |
| 20 | 17 | BC | BC | 71 |
| 21 | 18 | BC | BC | 63 |
| 22 | 19 | BC | BC | 59 |
| 23 | 19 | BC | BC | 50 |
| 24 | 20 | BC | BC | 39 |
| 25 | 21 | BC | BC | 30 |
| 26 | 22 | BC | BC | 56 |
| 27 | 23 | BC | BC | 49 |
| 28 | 24 | BOC | BC | 56 |
| 29 | 25 | BC | BC | 49 |
| 30 | 26 | BC | BC | 42 |
| 31 | 27 | BOC | BC | 48 |
| 32 | 28 | BOC | BC | 39 |
| 33 | 29 | BC | BC | 34/35 |
| 34 | 30 | BC | BC | 45 |
| 35 | 31 | BC | BC | 36 |
| 36 | 32 | BOC | BC | ? |
| 37 | 33 | BC | BC | 59 |
| 38 | 34 | OC | OC | ? |
| 39 | 35 | BC | BC | 47 |
| 40 | 36 | BC | BC | ? |
| 41 | 37 | BC | BC | 43 |
| 42 | 38 | BC | BC | 35 |
| 43 | 39 | BC | BC | 64 |

**BRCA1_2_PCR.HYB**

| | | | |
|---|---|---|---|
| **Stage 1:** | | | |
| Denaturation | 96°C | 8 min | |
| **Stage 2:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 63°C | 20 sec | 2 cycles |
| Extension | 72°C | 45 sec | |
| **Stage 3:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 61°C | 20 sec | 2 cycles |
| Extension | 72°C | 45 sec | |
| **Stage 4:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 59°C | 20 sec | 2 cycles |
| Extension | 72°C | 45 sec | |
| **Stage 5:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 58°C | 20 sec | 30 cycles |
| Extension | 72°C | 45 sec | |
| **Stage 6:** | | | |
| Hold | 12°C | | |

**BRCA1_2_PCR_20s.HYB**

| | | | |
|---|---|---|---|
| **Stage 1:** | | | |
| Denaturation | 95°C | 5 min | |
| **Stage 2:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 63°C | 20 sec | 2 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 3:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 61°C | 20 sec | 2 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 4:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 59°C | 20 sec | 2 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 5:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 58°C | 20 sec | 30 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 6:** | | | |
| Hold | 12°C | | |

**TD63-57_20s.HYB**

| | | | |
|---|---|---|---|
| **Stage 1:** | | | |
| Denaturation | 95°C | 5 min | |
| **Stage 2:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 63°C | 20 sec | 2 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 3:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 61°C | 20 sec | 2 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 4:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 59°C | 20 sec | 2 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 5:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 57°C | 20 sec | 30 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 6:** | | | |
| Hold | 12°C | | |

**TD63-57_45s.HYB**

| | | | |
|---|---|---|---|
| **Stage 1:** | | | |
| Denaturation | 95°C | 5 min | |
| **Stage 2:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 63°C | 20 sec | 2 cycles |
| Extension | 72°C | 45 sec | |
| **Stage 3:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 61°C | 20 sec | 2 cycles |
| Extension | 72°C | 45 sec | |
| **Stage 4:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 59°C | 20 sec | 2 cycles |
| Extension | 72°C | 45 sec | |
| **Stage 5:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 57°C | 20 sec | 30 cycles |
| Extension | 72°C | 45 sec | |
| **Stage 6:** | | | |
| Hold | 12°C | | |

**TD63-57_2min.HYB**

| | | | |
|---|---|---|---|
| **Stage 1:** | | | |
| Denaturation | 95°C | 5 min | |
| **Stage 2:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 63°C | 20 sec | 2 cycles |
| Extension | 72°C | 2 min | |
| **Stage 3:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 61°C | 20 sec | 2 cycles |
| Extension | 72°C | 2 min | |
| **Stage 4:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 59°C | 20 sec | 2 cycles |
| Extension | 72°C | 2 min | |
| **Stage 5:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 57°C | 20 sec | 30 cycles |
| Extension | 72°C | 2 min | |
| **Stage 6:** | | | |
| Hold | 12°C | | |

**Gradient_VUS_20s.HYB**

| | | | |
|---|---|---|---|
| **Stage 1:** | | | |
| Denaturation | 95°C | 5 min | |
| **Stage 2:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 55°C – Gradient: 15 | 20 sec | 30 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 3:** | | | |
| Final extension | 72°C | 8 min | |
| **Stage 4:** | | | |
| Hold | 12°C | | |

**VUS_58.HYB**

| | | | |
|---|---|---|---|
| **Stage 1:** | | | |
| Denaturation | 95°C | 5 min | |
| **Stage 2:** | | | |
| Denaturation | 95°C | 20 sec | |
| Annealing | 58°C | 20 sec | 36 cycles |
| Extension | 72°C | 20 sec | |
| **Stage 3:** | | | |
| Hold | 12°C | | |

**PALB2_4C.HYB**

| | | | |
|---|---|---|---|
| **Stage 1:** | | | |
| Denaturation | 95°C | 5 min | |
| **Stage 2:** | | | |
| Denaturation | 95°C | 20 sec | 36 cycles |
| Annealing | 65°C | 20 sec | |
| Extension | 72°C | 20 sec | |
| **Stage 3:** | | | |
| Hold | 12°C | | |

**Cycle parameters for APRT-test:**

| | | | |
|---|---|---|---|
| **Stage 1:** | | | |
| Denaturation | 94°C | 3 min | |
| **Stage 2:** | | | |
| Denaturation | 94°C | 30 sec | 30 cycles |
| Annealing | 63°C | 30 sec | |
| Extension | 72°C | 1 min | |
| **Stage 3:** | | | |
| Final extension | 72°C | 7 min | |
| **Stage 4:** | | | |
| Hold | 12°C | | |

# Appendix B – Primers, size markers and reagents

**Table 20 Specific PALB2 primers used in this study. Primers are from the following article: Fanconi anemia is associated with a defect in the BRCA2 partner PALB2 [105], except for PALB2ex4C.2R (in bold), which was made with primer3. The corresponding primer from the article was located in the end of exon 4. Sequence marked in bold corresponds to the forward and revers M13 sequence in common for all these primers.**

| Exon | Name | Primer sequence (5'-3') | Amplicon size (bp) |
|---|---|---|---|
| **Exon 1** | PALB2ex1.F | **TGTAAAACGACGGCCAGT**ACAGCGCGGCTCTCCTTTAG | 410 |
| | PALB2ex1.R | **CAGGAAACAGCTATGACC**ATACTGCTGCCCTCGGACTG | |
| **Exon 2,3** | PALB2ex2,3.F | **TGTAAAACGACGGCCAGT**GTAGATTGTTATGGACCAGTGCTACT | 511 |
| | PALB2ex2,3.R | **CAGGAAACAGCTATGACC**GTCTATTGCTAGTCATTATCTTCACAC | |
| **Exon 4A** | PALB2ex4A.F | **TGTAAAACGACGGCCAGT**TCTGCCTGAATGAAATGTCACTGA | 645 |
| | PALB2ex4A.R | **CAGGAAACAGCTATGACC**GGTCTTCTTAGGAATGTATCAACACC | |
| **Exon 4B** | PALB2ex4B.F | **TGTAAAACGACGGCCAGT**CAGATTCTCCAGAACCAGTTACAGAA | 559 |
| | PALB2ex4B.R | **CAGGAAACAGCTATGACC**TTTTCTGCAGAAAGAGGAGAGGTT | |
| **Exon 4C** | PALB2ex4C.F | **TGTAAAACGACGGCCAGT**CTCCCAGTGACACTCTTGATGGC | 734 |
| | **PALB2ex4C.2R** | **CAGGAAACAGCTATGACC**ATTACAGACGTAAGCCACCACAC | |
| **Exon 5A** | PALB2ex5A.F | **TGTAAAACGACGGCCAGT**TTAAATCTAGGAGATCCTATTCTCTTTG | 531 |
| | PALB2ex5A.R | **CAGGAAACAGCTATGACC**GTATAAAGTAATATGGATGAAGAAAGGC | |
| **Exon 5C** | PALB2ex5C.F | **TGTAAAACGACGGCCAGT**AGAAATGGAGGACTTAGAAGAGGAC | 586 |
| | PALB2ex5C.R | **CAGGAAACAGCTATGACC**CATGCTGTTTACATTCACTAAGGC | |
| **Exon 6** | PALB2ex6.F | **TGTAAAACGACGGCCAGT**GCTGCTGTTATAAGAGGAAATAAAGACA | 344 |
| | PALB2ex6.R | **CAGGAAACAGCTATGACC**GGGAAAAATAACCAATCCAAATCTG | |
| **Exon 7** | PALB2ex7.F | **TGTAAAACGACGGCCAGT**GCTCTTTCTTTTCACCTGCATAAGA | 492 |
| | PALB2ex7.R | **CAGGAAACAGCTATGACC**TGGGTATATGGGTGCTCACTATACA | |
| **Exon 8** | PALB2ex8.F | **TGTAAAACGACGGCCAGT**CCTTGTACAGTGAGAATACAAAAGAATGTGA | 270 |
| | PALB2ex8.R | **CAGGAAACAGCTATGACC**TAGGTTATTACCTGCACTTAAAACCA | |
| **Exon 9** | PALB2ex9.F | **TGTAAAACGACGGCCAGT**AAAAAAGTGAACCTAGTCCTTTAATATT | 581 |
| | PALB2ex9.R | **CAGGAAACAGCTATGACC**GCTCAAACTTCTGCCTTGGC | |
| **Exon 10** | PALB2ex10.F | **TGTAAAACGACGGCCAGT**ATATTATGCAGTTCAACAATGCGG | 386 |
| | PALB2ex10.R | **CAGGAAACAGCTATGACC**ACCTGGGTGATAGGAGGAGACTC | |
| **Exon 11** | PALB2ex11.F | **TGTAAAACGACGGCCAGT**ACCTCCTAAGACATGCTATGATGAATAA | 379 |
| | PALB2ex11.R | **CAGGAAACAGCTATGACC**GCCAGAAAATTTACCAAGCAATCA | |
| **Exon 12** | PALB2ex12.F | **TGTAAAACGACGGCCAGT**CAGAGCCTATCGGTCATTGCTT | 438 |
| | PALB2ex12.R | **CAGGAAACAGCTATGACC**TCTGGGGTTTGACTCAAGTCCA | |
| **Exon 13** | PALB2ex13.F | **TGTAAAACGACGGCCAGT**TTTAATTGTTTTTTGGATATGTAATCTGAA | 549 |
| | PALB2ex13.R | **CAGGAAACAGCTATGACC**AAATATTTATTGCCATTTGAAGCTTTAT | |

**Table 21. Primers used for the variants in *BRCA1* and *BRCA2* of unknown clinical significance. Complementary DNA (cDNA) and genomic DNA (gDNA) is listed to show that we have taken the presence of gDNA into consideration during primer design. Primer design was done with Primer3.**

| Exon | Name | Primer sequence (5'-3') | Amplicon size (bp) cDNA/gDNA | PCR-program |
|---|---|---|---|---|
| **1**<br>**3** | BRCA1 del_intron1.F<br>BRCA1 del_intron1.R | CTCGCTGAGACTTCCTGGAC<br>TGTGGAGACAGGTTCCTTGA | 228/ 1.482 | TD63-57_20s.HYB |
| **2**<br>**7** | BRCA1 130.F<br>BRCA1 130.R | GCTCTTCGCGTTGAAGAAGT<br>GAAGTCTTTTGGCACGGTTT | 400/ 19.934 | VUS_58.HYB |
| **7**<br>**11** | BRCA1 486.F<br>BRCA1 486.R | CATCCAAAGTATGGGCTACAGA<br>GCTGTAATGAGCTGGCATGA | 482/ 9.514 | TD63-57_20s.HYB |
| **10**<br>**11** | BRCA1_1419.2F<br>BRCA1_1419.R | GAGGACAAAGCAGCGGATAC<br>CCGTTTGGTTAGTTCCCTGA | 1.199/ 5.990 | TD63-57_45s.HYB |
| **11**<br>**13** | BRCA1 2521.F<br>BRCA1 2521.2R | CAGCATTTGAAAACCCCAAG<br>AGTTCAGCCATTTCCTGCTG | 1.825/ 10.589 | TD63-57_2min.HYB |
| **11**<br>**13** | BRCA1 3418.F<br>BRCA1 2521.2R | TAGGGGTTTTGCAACCTGAG<br>AGTTCAGCCATTTCCTGCTG | 985/ 9.755 | TD63-57_45s.HYB |
| **16**<br>**20** | BRCA1 5075.F<br>BRCA1 5075.R | GGGAGAAGCCAGAATTGACA<br>TGGACCTTGGTGGTTTCTTC | 347/ 13.932 | TD63-57_20s.HYB |
| **21**<br>**3'UTR** | BRCA1 5434.F<br>BRCA1 5434.R | TTCAGGGGGCTAGAAATCTG<br>AAGCTCATTCTTGGGGTCCT | 360/ 5.485 | VUS_58.HYB |
| **1**<br>**4** | BRCA2 e1F-ny<br>BRCA2 e4R-ny | AGCGTGAGGGGACAGATTTG<br>GTGGACAGGAAACATCATCTGC | 519/ 9.572 | TD63-57_20s.HYB |
| **7**<br>**10** | BRCA2 e7.F<br>BRCA2 e10.R | AGGAGCTGAGGTGGATCCTG<br>TTTCCAATGTGGTCTTTGCAG | 303/ 5.799 | TD63-57_20s.HYB |
| **10**<br>**11** | BRCA2 e10.F<br>BRCA2 2680.R | GTTCAGCCCAGTTTGAAGCA<br>TGACACTTGGGTTGCTTGTT | 980/ 3.857 | TD63-57_45s.HYB |
| **10**<br>**11** | BRCA2 e10.F<br>BRCA2 4068.R | GTTCAGCCCAGTTTGAAGCA<br>CTTGAGCTTTCGCAACTTCC | 2.343/ 5.220 | TD63-57_2min.HYB |
| **11**<br>**13** | BRCA2 4828.F<br>BRCA2 e13.R | CAATGGGCAAAGACCCTAAA<br>GTGCCATCTGGAGTGCTTTT | 2.324/ 7.858 | TD63-57_2min.HYB |
| **11**<br>**13** | BRCA2 5272_5284delAAT.F<br>BRCA2 e13.R | TTTGATGGTCAACCAGAAAGAA<br>GTGCCATCTGGAGTGCTTTT | 1.916/ 7.450 | TD63-57_2min.HYB |
| **11**<br>**13** | BRCA2 6100.F<br>BRCA2 e13.R | CGCAAGACAAGTGTTTTCTGA<br>GTGCCATCTGGAGTGCTTTT | 1.023/ 6.557 | TD63-57_45s.HYB |
| **13**<br>**15** | BRCA2 7301.F<br>BRCA2 7301.R | CCGATTACCTGTGTACCCTTTC<br>GAGGGAACTTGGCCTCCTAC | 614/ 9.717 | TD63-57_20s.HYB |
| **16**<br>**18** | BRCA2 7878.F<br>BRCA2 7878.R | TACAGTTGGCTGATGGTGGA<br>GCTGTGTCATCCCTTTCCAT | 310/ 5.388 | TD63-57_20s.HYB |
| **17**<br>**18** | BRCA2 8177.F<br>BRCA2 8177.R | ATGGAAACTGGCAGCTATGG<br>CGATGATAAGGGCAGAGGAA | 538/ 7.891 | TD63-57_20s.HYB |
| **22**<br>**25** | BRCA2 9116.F<br>BRCA2 9116.F | TAGGAAGGCCATGGAATCTG<br>GCCTGATTTGGATTCTGGTC | 559/ 15.429 | TD63-57_20s.HYB |

**Table 22 Primers used for sequencing the entire *BRCA1* insert in plasmid. Primer sequences in bold are reverse primers. The first and last primers are located in plasmid sequence.**

| Name | Primer sequence (5'-3') |
|---|---|
| **T7 primer** | TAATACGACTCACTATAGGG |
| **BRCA1 130.F** | GCTCTTCGCGTTGAAGAAGT |
| **BRCA1 130.R** | **GAAGTCTTTTGGCACGGTTT** |
| **BRCA1 486.F** | CATCCAAAGTATGGGCTACAGA |
| **BRCA1 1419.F** | CTCAAGGAACCAGGGATGAA |
| **BRCA1ex11_fr2.F** | CAGAATGAATGTAGAAAAGGCTGA |
| **BRCA1ex11_fr3.F** | TGAAAGAGTTCACTCCAAATCAG |
| **BRCA1ex11_fr4.F** | TGCTTTCAAAACGAAAGCTG |
| **BRCA1ex11_fr5.F** | TTTGTCAATCCTAGCCTTCCA |
| **BRCA1ex11._fr6.F2** | CATTGGGACATGAAGTTAACCA |
| **BRCA1ex11_fr7.F** | CAGTGAGCACAATTAGCCGTA |
| **BRCA1ex11.2F3** | TGTTCTGAGACACCTGATGACC |
| **BRCA1 5075.F** | GGGAGAAGCCAGAATTGACA |
| **BRCA1_2521.2R** | **AGTTCAGCCATTTCCTGCTG** |
| **BRCA1 5434.F** | TTCAGGGGGCTAGAAATCTG |
| **BGH reverse** | TAGAAGGCACAGTCGAGGC |

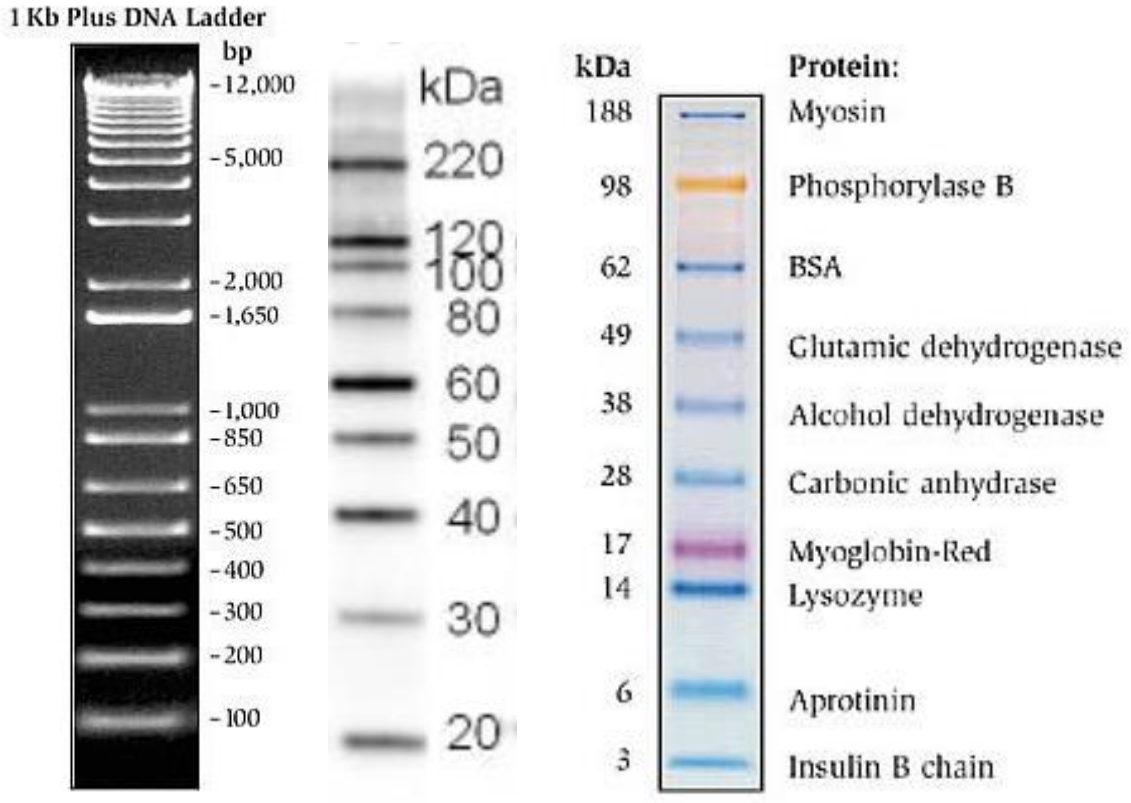



**Figure 32. Size markers for gel-electrophoresis and western blot: 1 Kb Plus DNA ladder (Invitrogen); MagicMark™XP Western Protein Standard loaded on a NuPAGE®Novex® 4–12% Bis-Tris Gel (Life Technologies); Apparent molecular weights of SeeBlue Plus2 Pre-Stained Standard on a NuPAGE®Novex® 4–12% Bis-Tris Gel (Life Technologies).**

# Buffers and cell medium

**PBS:**
1.4 mM $KH_2PO_4$
8 mM $Na_2HPO_4$
140 mM NaCl
2.7 mM KCl, pH 7.3

**5xTBE:**
54g Tris base
27.5g Boric acid
20 ml of 0.5M EDTA (pH 8.0)

**Transfer buffer (Western Blot):**
50ml NuPAGE transfer buffer
1 ml NuPAGE antioxidant
100 ml methanol
1000 ml MilliQ water

**Blocking buffer (Western Blot):**
100 ml PBS
5 g dry low fat milk
100 µl Tween 20 (purchased from Sigma-Aldrich)

**Growth medium:**
500 ml Dulbecco's Modified Eagle Medium (Gibco® Life Technologies)
50ml fetal bovine serum
3 ml penicillin (10.000 units/ml) Streptomycin (10.000 µg/ml) (Gibco® Life Technologies)

**1L LB medium(BD Difco Lauria -Bertani Broth Miller REF 244620):**
10 g trytone
5 g yeast extract
10 g NaCl
Bring total volume up to 1L with MilliQ water

**LB Agar (BD Difco Lauria -Bertani Broth Agar REF 244520):**
10 g trytone
5 g yeast extract
10 g NaCl
15g Agar
Bring total volume up to 1L with MilliQ water

# Appendix C – results

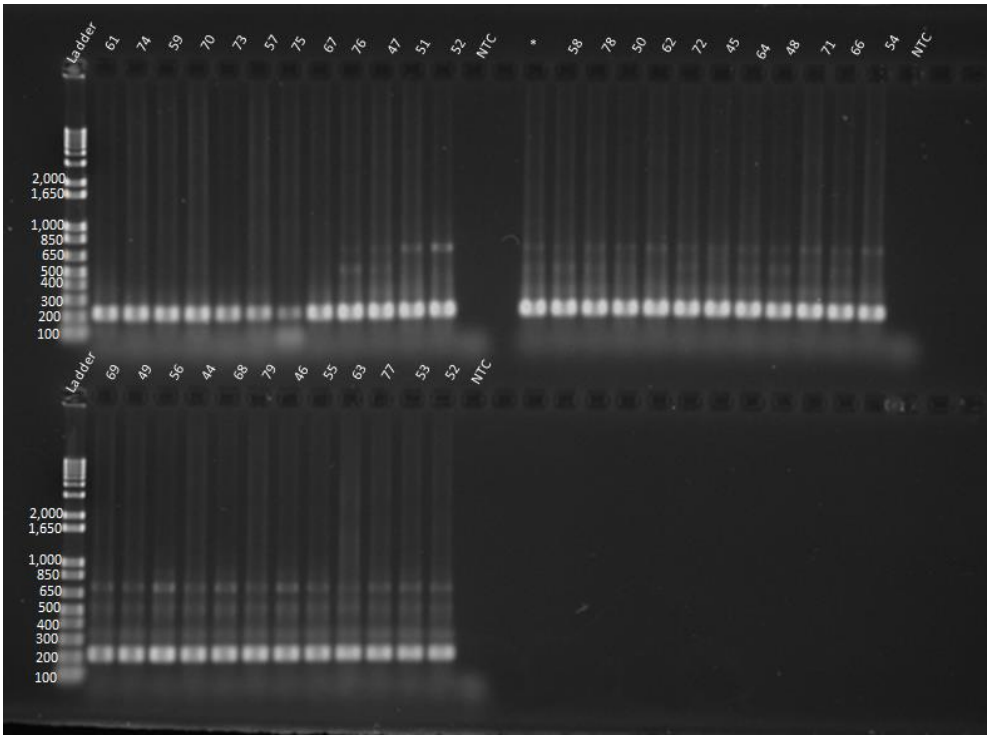



**Figure 33. Gel image from APRT-test on patient samples. The ladder is 1 Kb plus DNA molecular marker (Invitrogen). Sample number is listed at the top and only missing 120483 (normal control) and sample 65. Sample marked "*" was not used in this study.**

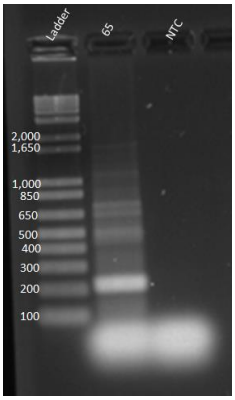



**Figure 34. Gel image from APRT-test on patient sample 65 together with 120483 (normal control). The ladder is 1 Kb plus DNA molecular marker (Invitrogen).**

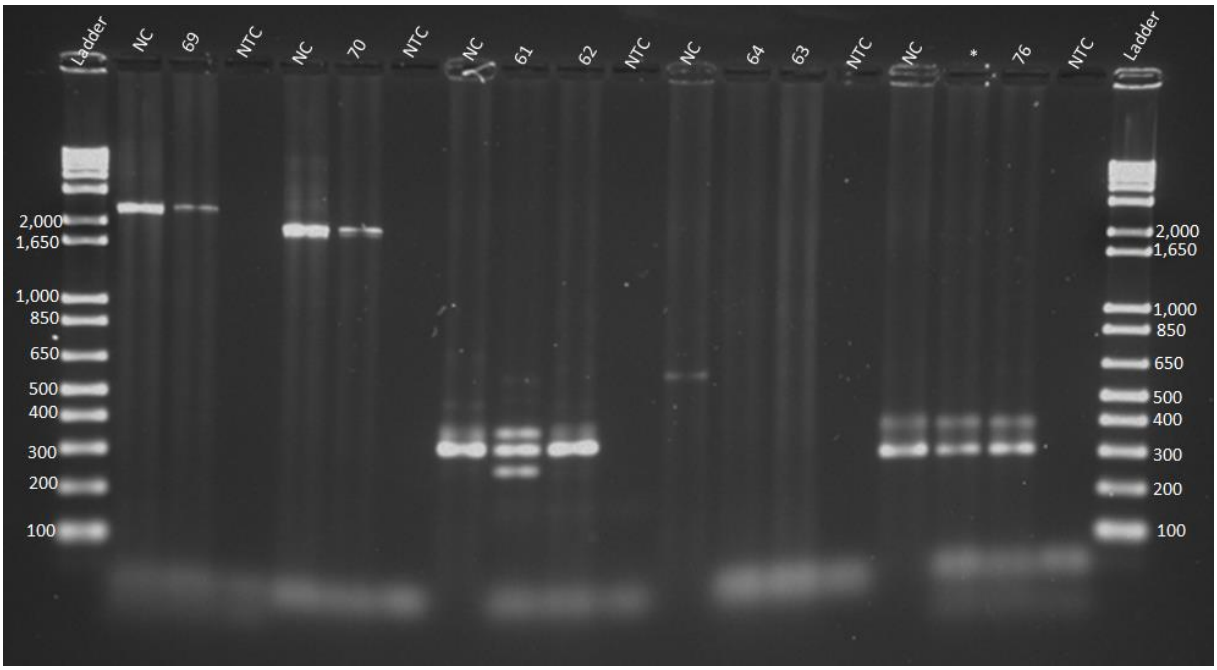



**Figure 35. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen). Samples are listed at the top, together with normal control (NC) and no template control (NTC). Well marked "*" was amplification of wrong sample and had to be redone. Samples 63 and 64 together with corresponding NC was later amplified with new primers.**

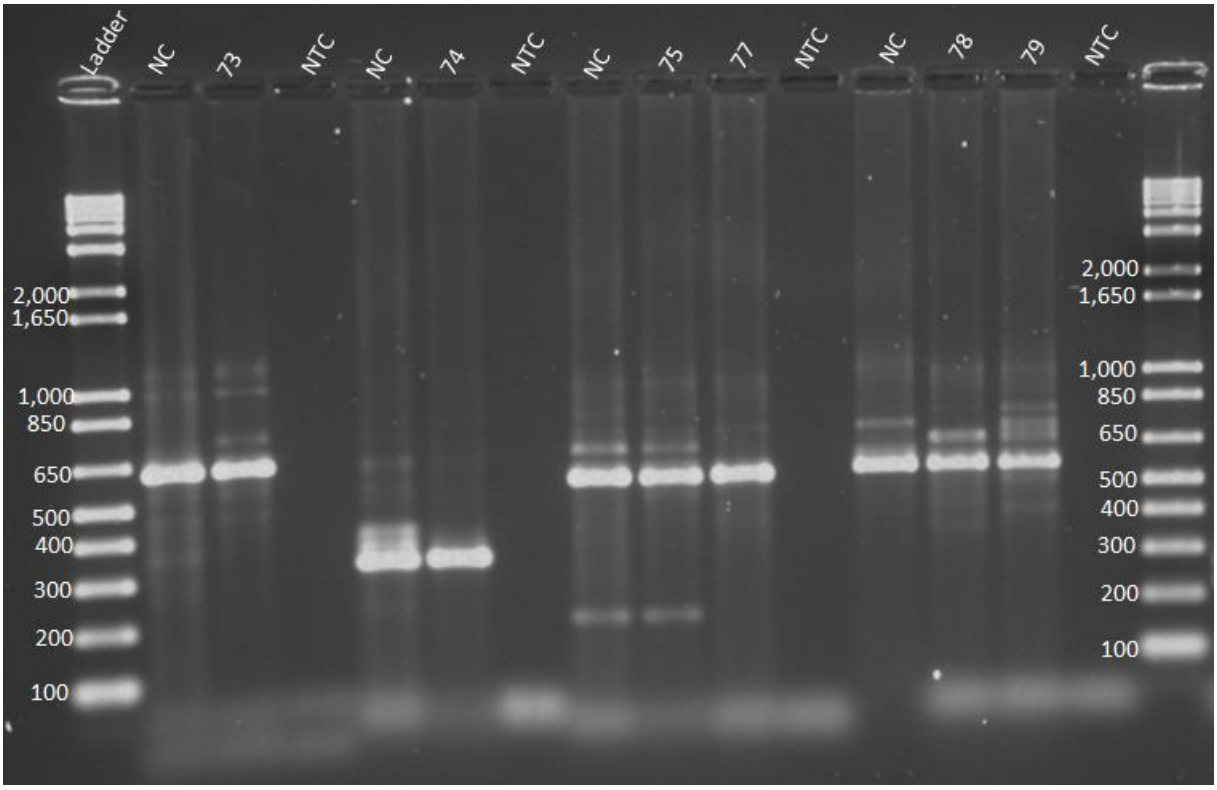



**Figure 36. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen). Samples are listed at the top, together with normal control (NC) and no template control (NTC).**

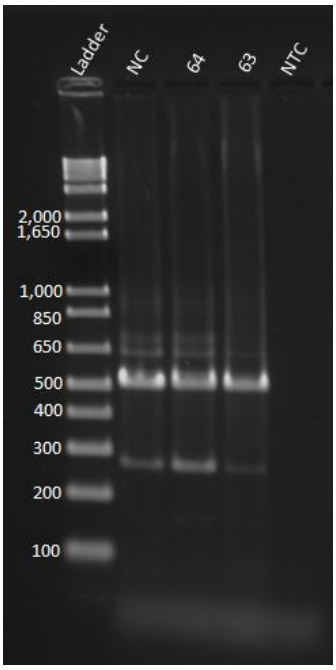



**Figure 37. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen). Samples are listed at the top, together with normal control (NC) and no template control (NTC).**

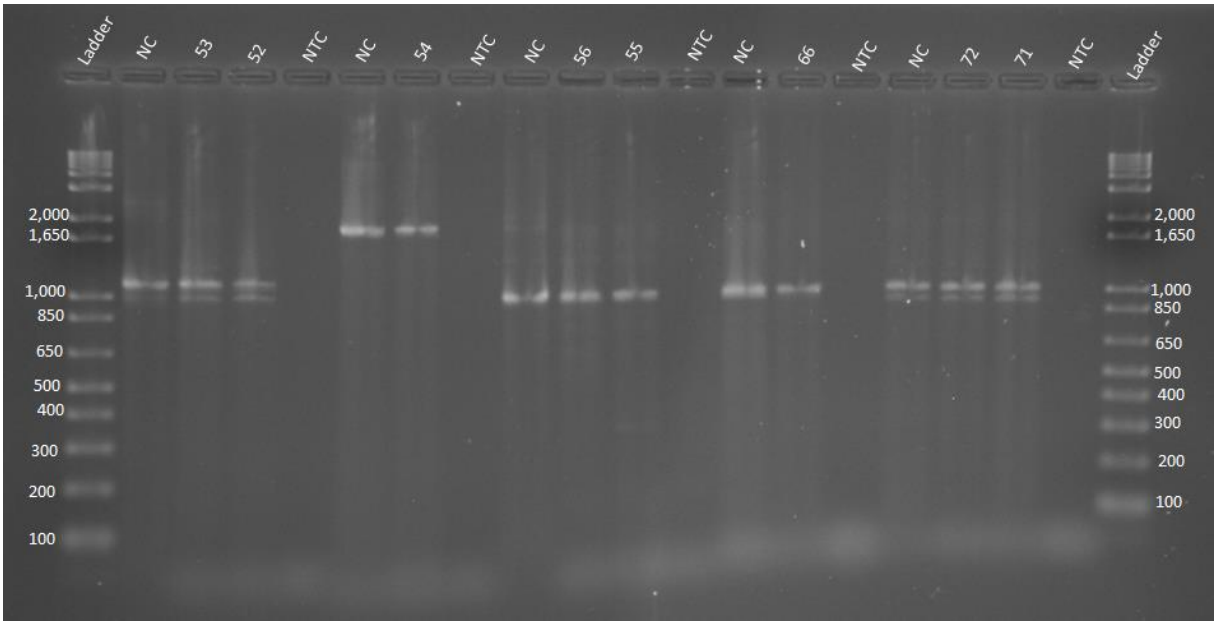



**Figure 38. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen). Samples are listed at the top, together with normal control (NC) and no template control (NTC).**

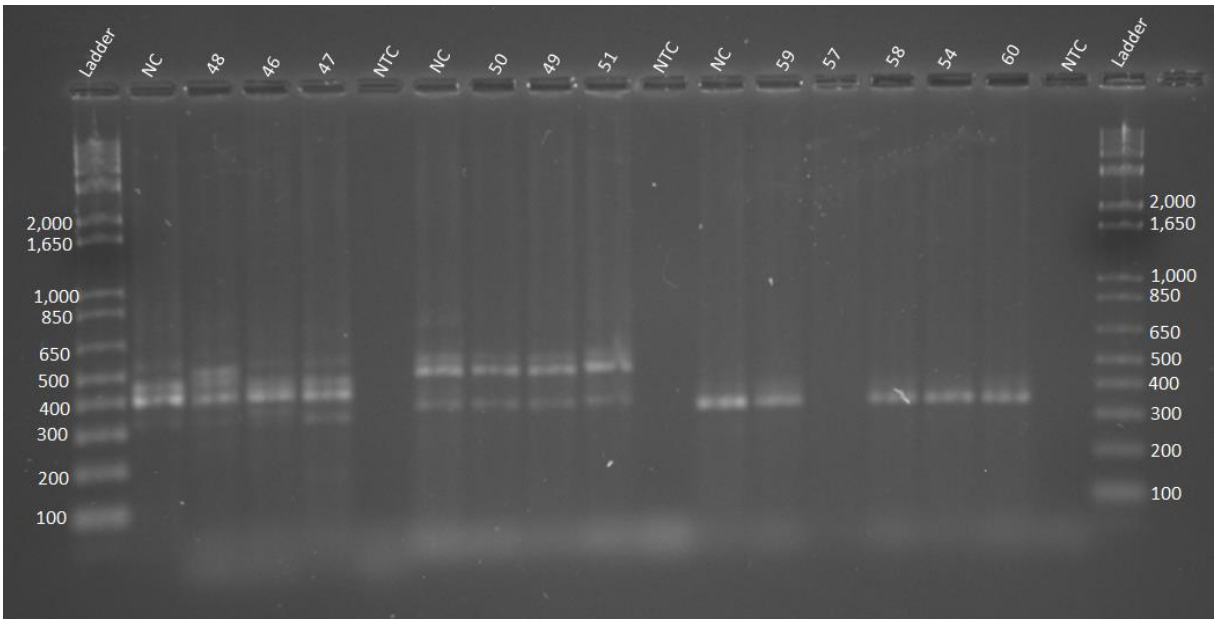



**Figure 39. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen). Samples are listed at the top, together with normal control (NC) and no template control (NTC). Sample 57 was later re-amplified.**

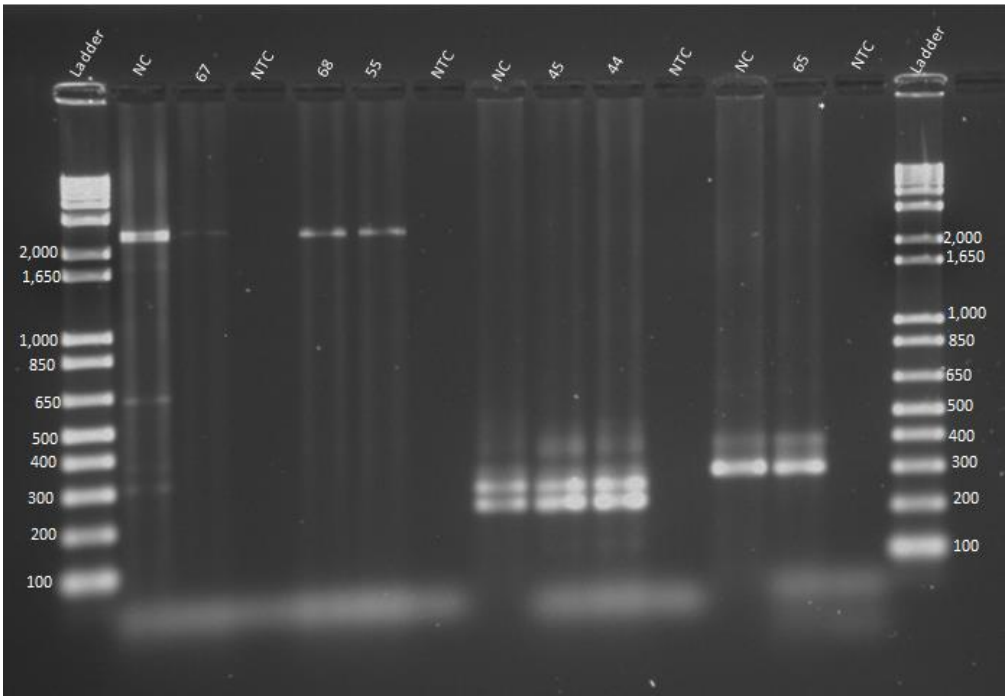



**Figure 40. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen). Samples are listed at the top, together with normal control (NC) and no template control (NTC). Sample 67, 68 and 55 share a common NC. The three samples were not sufficient to produce readable sequences.**

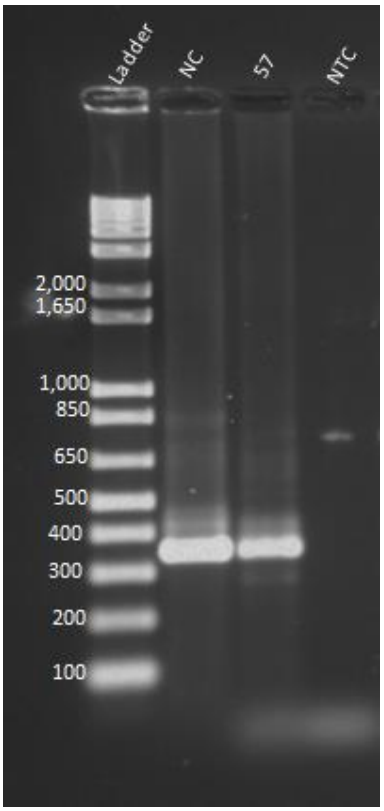



**Figure 41. PCR products amplified from different cDNA samples were run on 2% agarose gel. Ladder is 1kB plus molecular weight ladder (Invitrogen). Sample 57 is marked at the top, together with normal control (NC) and no template control (NTC).**

**Table 23. Overview of the variants found in *BRCA1* together with results from prediction programs. Sorting Intolerant From Tolerant (SIFT) and Align Grantham Variation, Grantham Deviation (AlignGVGD) are both prediction programs based on the evolutionary conservation at their location within the protein, while AlignGVGD also takes into consideration amino acid characteristics. Polyphen2 is divided in HumDiv and HumVar, difference being the training sets. MutationTaster prediction is based on evolutionary conservation, splice-site changes, loss of protein features and changes that might affect amount of RNA. Classification is based on both the prediction programs and mutation database and literature searches. Comment on splicing is based on corresponding prediction programs in Alamut v.2.2 rev.0, "weaker" or "stronger" is according to original splice predictions at original site splice site unless otherwise mentioned. Nomenclature is according to HGVS standards, c.1 corresponds to the first nucleotide in ATG.**

| Mutation | Localization | Protein | SIFT | AlignGVGD | PolyPhen | | MutationTaster | Comment on splicing prediction | class |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | HumDiv | HumVar | | | |
| **c.-33-29delAAAAA** | Intron 1 | - | - | - | - | - | - | Weaker 3'ss downstream | 3 |
| **c.1A>G** | Exon 2 | p.Met1? | - | - | - | - | - | | 5 |
| **c.130T>A** | Exon 3 | p.Cys44Ser | Deleterious (0.00) | C65 | Probably damaging (0.979) | Possibly damaging (0.760) | Disease causing (0.995) | | 5 |
| **c.140G>T** | Exon 5 | p.Cys47Phe | Deleterious (0.00) | C65 | Possibly damaging (0.476) | Benign (0.145) | Disease causing (0.991) | | 4 |
| **c.213-5T>A** | Intron 5 | - | - | - | - | - | - | Weaker 3'ss downstream | 4 |
| **c.486G>T** | Exon 8 | p.= (p.Val162Val) | - | - | - | - | - | Indication of new 3'ss (HSF and MES) | 2 |
| **c.548-17G>T** | Intron 8 | - | - | - | - | - | - | Stronger 3'ss downstream | 1 |
| **c.734A>T** | Exon 11 | p.Asp245Val | Deleterious (0.00) | C65 | Probably damaging (0.979) | Possibly damaging (0.760) | Disease causing (0.995) | | 3 |
| **c.1419C>T** | Exon 11 | p.= (p.Asn473Asn) | - | - | - | - | - | ~750 bp into exon 11, indication of new 3'ss (MES) | 2 |
| **c.1487G>A** | Exon 11 | p.Arg496His | Tolerated (0.23) | C0 | Benign (0.001) | Benign (0.004) | Polymorphism (1.0) | | 1 |
| **c.2521C>T** | Exon 11 | p.Arg841Trp | Deleterious (0.02) | C15 | Benign (0.001) | Benign (0.006) | Polymorphism (1.0) | | 1 |
| **c.3418A>G** | Exon 11 | p.Ser1140Gly | Tolerated (0.43) | C0 | Benign (0.002) | Benign (0.008) | Polymorphism (1.0) | | 1 |
| **c.3708T>G** | Exon 11 | p.Asn1236Lys | Tolerated (0.55) | C0 | Benign (0.031) | Benign (0.006) | Polymorphism (1.0) | | 3 |
| **c.5075A>C** | Exon 18 | p.Asp1692Ala | Deleterious (0.00) | C65 | Disease causing (1.0) | Benign (0.017) | Benign (0.003) | Possibly a weaker 3'ss (variant affects the first nucleotide in exon18) | 3 |
| **c.5096G>A** | Exon 18 | p.Arg1699Gln | Deleterious (0.00) | C35 | Probably damaging (0.999) | Probably damaging (0.915) | Disease causing (0.999) | Possibly new alternative 3'ss 24 nucleotides into exon 18 | 5 |
| **c.5117G>C** | Exon 18 | p.Gly1706Ala | Deleterious (0.00) | C55 | Possibly damaging (0.907) | Benign (0.415) | Disease causing (1.0) | | 1 |
| **c.5123C>T** | Exon 18 | p.Ala1708Val | Deleterious (0.00) | C65 | Probably damaging (1.000) | Possibly damaging (0.686) | Disease causing (1.0) | GeneSplicer predicts weaker 5'ss downstream. | 4 |
| **c.5125G>A** | Exon 18 | p.Gly1709Arg | Deleterious (0.01) | C15 | Possibly damaging (0.788) | Benign (0.317) | Disease causing (1.0) | | 3 |
| **c.5434C>G** | Exon 23 | p.Pro1812Ala | Tolerated (0.05) | C0 | Benign (0.016) | Benign (0.001) | Disease causing (0.961) | Stronger cryptic splice site at mutation site | 5 |
| **c.5513T>G** | Exon 24 | p.Val1838Gly | Deleterious (0.00) | C35 | Probably damaging (1.000) | Probably damaging (0.998) | Disease causing (1.0) | | 3 |

**Table 24 Overview of the variants found in *BRCA2* together with results from prediction programs. Sorting Intolerant From Tolerant (SIFT) and Align Grantham Variation, Grantham Deviation (AlignGVGD) are both prediction programs based on the evolutionary conservation at their location within the protein, while AlignGVGD also takes into consideration amino acid characteristics. Polyphen2 is divided in HumDiv and HumVar, difference being the training sets. MutationTaster prediction is based on evolutionary conservation, splice-site changes, loss of protein features and changes that might affect amount of RNA. Classification is based on both the prediction programs and mutation database and literature searches. Comment on splicing is based on corresponding prediction programs in Alamut v.2.2 rev.0, "weaker" or "stronger" is according to original splice predictions at original site splice site unless otherwise mentioned. Nomenclature is according to HGVS standards, c.1 corresponds to the first nucleotide in ATG**

| Mutation | Localization | Protein | SIFT | AlignGVGD | PolyPhen | | MutationTaster | Comment on splicing | Class. |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | HumDiv | HumVar | | | |
| **c.40A>G** | Exon 2 | p.Ile14Val | Tolerated (0.92) | C0 | Benign (0.227) | Benign (0.034) | Polymorphism (0.996) | Possible new 3'ss at mutation site | 2 |
| c.68-7T>A | **Intron 2** | - | - | - | - | - | - | **Weaker 3'ss** | **3** |
| **c.750G>A** | Exon 9 | p.= (p.Val250Val) | - | - | - | - | - | Possible stronger cryptic 3'ss 5 nucleotides downstream of mutation | 3 |
| **c.2680G>A** | Exon 11 | p.Val894Ile | Tolerated (0.53) | C0 | Benign (0.001) | Benign (0.001) | Polymorphism (1.0) | Possible new 3'ss 4 nucleotides downstream of mutation | 1 |
| **c.3568C>T** | Exon 11 | p.Arg1190Trp | Deleterious (0.01) | C25 | Possibly damaging (0.876) | Benign (0.153) | Polymorphism (1.0) | | 1 |
| **c.4068G>A** | Exon 11 | p.= (p.Leu1356Leu) | - | - | - | - | - | | 3 |
| **c.4828G>A** | Exon 11 | p.Val1610Met | Tolerated (0.10) | C0 | Benign (0.019) | Benign (0.006) | Polymorphism (1.0) | | 2 |
| **c.5272_5274delAAT** | Exon 11 | p.Asn1758del | - | - | - | - | - | | 2 |
| **c.6100C>T** | Exon 11 | p.Arg2034Cys | Tolerated (0.07) | C0 | Possibly damaging (0.876) | Benign (0.118) | Polymorphism (1.0) | Stronger cryptic 3'ss 10 nucleotides downstream of mutation | 1 |
| **c.6821G>T** | Exon 11 | p.Gly2274Val | Tolerated (0.24) | C0 | Probably damaging (0.994) | Possibly damaging (0.748) | Disease causing (1.0) | | 2 |
| **c.7301A>C** | Exon 14 | p.Lys2434Thr | Tolerated (0.20) | C0 | Benign (0.004) | Benign (0.003) | Polymorphism (1.0) | | 3 |
| **c.7878G>C** | Exon 17 | p.Trp2626Cys | Deleterious (0.00) | C65 | Probably damaging (1.000) | Probably damaging (1.000) | Disease causing (1.0) | | 5 |
| **c.8177A>G** | Exon 18 | p.Tyr2726Cys | Deleterious (0.00) | C65 | Probably damaging (1.000) | Probably damaging (1.000) | Disease causing (1.0) | | 4 |
| **c.8323A>G** | Exon 18 | p.Met2775Val | Tolerated (0.23) | C0 | Possibly damaging (0.956) | Possibly damaging (0.715) | Disease causing (0.893) | | 3 |
| **c.9116C>T** | Exon 23 | p.Pro3039Leu | Deleterious (0.02) | C0 | Possibly damaging (0.579) | Benign (0.032) | Disease causing (1.0) | Weaker 5'ss downstream (variant is 2 nucleotides from original 5'ss in exon 23) | 3 |

**Table 25. Protein concentration after transfection with fugene and cell harvesting using M-PER. Negative controls are lysate from untreated cells. The fugene controls are lysate from cells only treated with the transfection agent fugene. pcDNA 3.1 was only transfected with the corresponding empty plasmid. BRCA1 was transfected with the HA-pcDNA BRCA1 Ser1387Ala.**

| Plate number | Sample name | Concentration (mg/ml) | cell type |
|---|---|---|---|
| **1** | negative control | 0.865 | A375 |
| **1** | Fugene control | 0.749 | A375 |
| **1** | pcDNA 3.1, 24h | 0.223 | A375 |
| **1** | pcDNA 3.1, 48h | 1.266 | A375 |
| **1** | BRCA1, 24h | 0.739 | A375 |
| **1** | BRCA1, 48h | 0.813 | A375 |
| **2** | negative control | 0.631 | HEK-293 |
| **2** | Fugene control | 0.634 | HEK-293 |
| **2** | pcDNA 3.1, 24h | 1.220 | HEK-293 |
| **2** | pcDNA 3.1, 48h | 1.111 | HEK-293 |
| **2** | BRCA1, 24h | 0.742 | HEK-293 |
| **2** | BRCA1, 48h | 1.102 | HEK-293 |